OM nucleic - nucleic search, using sw model

Run on: March 9, 2002, 01:26:12 ; Search time 156.11 Seconds
(without alignments)
6343.032 Million cell updates/sec

Title: US-09-699-295-301
Perfect score: 1155
Sequence: 1 atggtgattgaggttgattc ......caagaacagaaataaataa 1155

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_1101:*
1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1155	100.0	1155	21	AAC81011	Human B11Aq1 antig
	2	1155	100.0	1155	21	AAA06598	Human immunogenic
	3	1155	100.0	1155	22	AAH93714	Human prostate spe
	4	1155	100.0	1155	22	AAH85028	Human prostate spe
	5	1155	100.0	1155	22	AAH02779	Prostate tumour an
	6	1131.6	98.0	2000	21	AAC81012	Human B11Aq1 antig
	7	1131.6	98.0	2000	22	AAH93715	Human prostate spe
	8	1131.6	98.0	2000	22	AAH85029	Human prostate spe
	9	1131.6	98.0	2000	22	AAH02780	Prostate tumour an
	10	1131.6	98.0	2040	21	AAC81013	Human B11Aq1 antig
	11	1131.6	98.0	2040	21	AAA06600	Human immunogenic
	12	1131.6	98.0	2040	22	AAH93716	Human prostate spe
	13	1131.6	98.0	2040	22	AAH85030	Human prostate spe
	14	1131.6	98.0	2040	22	AAH02781	Prostate tumour an
	15	1130	97.8	2000	21	AAA06599	Human immunogenic
	16	1088.6	94.3	1853	19	AAV68996	DNA molecule encod
	17	1088.6	94.3	1853	21	AAC81007	Human B11Aq1 antig
	18	1088.6	94.3	1853	21	AAA06594	Human immunogenic
	19	1088.6	94.3	1853	22	AAH93710	Human prostate spe
	20	1088.6	94.3	1853	22	AAH85024	Human prostate spe
	21	1088.6	94.3	1853	22	AAH02775	Prostate tumour an
	22	841	72.8	1852	21	AAC81015	Human breast tumou
	23	841	72.8	1852	22	AAH93825	Human prostate spe
	24	841	72.8	1852	22	AAH85139	Human prostate spe
c	25	829	71.8	1851	19	AAV68992	DNA molecule encod
c	26	829	71.8	1851	19	AAV68993	DNA molecule encod
c	27	829	71.8	1851	21	AAC81003	Human breast tumou
c	28	829	71.8	1851	21	AAC81004	Human B11Aq1 antig
c	29	829	71.8	1851	21	AAA06591	Human immunogenic
c	30	829	71.8	1851	22	AAH93707	Human prostate spe
c	31	829	71.8	1851	22	AAH85021	Human prostate spe
c	32	829	71.8	1851	22	AAH02772	Prostate tumour an
	33	829	71.8	2184	19	AAV68997	DNA molecule encod
	34	829	71.8	2184	21	AAC81008	Human B11Aq1 antig
	35	829	71.8	2184	21	AAA06595	Human immunogenic
	36	829	71.8	2184	22	AAH93711	Human prostate spe
	37	829	71.8	2184	22	AAH85025	Human prostate spe
	38	829	71.8	2184	22	AAH02776	Prostate tumour an
	39	804.8	69.7	1512	19	AAV68995	DNA molecule encod
	40	804.8	69.7	1512	21	AAC81006	Human B11Aq1 antig
	41	804.8	69.7	1512	21	AAA06593	Human immunogenic
	42	804.8	69.7	1512	22	AAH93709	Human prostate spe
	43	804.8	69.7	1512	22	AAH85023	Human prostate spe
	44	804.8	69.7	1512	22	AAH02774	Prostate tumour an
	45	768	66.5	879	21	AAC81016	Human breast tumou

ALIGNMENTS

RESULT 1
AAC81011
ID AAC81011 standard; cDNA; 1155 BP.
XX
AC AAC81011;
XX
DT 13-FEB-2001 (first entry)
XX
DE Human B11Aq1 antigen splice isoform B11-15 cDNA.
XX
KW Human; breast tumour-specific antigen; cytostatic; vaccine;
KW breast cancer; B18Aq1; B11Aq1; B15Aq1; ss.
XX
OS Homo sapiens.
XX
PN WO200061753-A2.
XX
PD 19-OCT-2000.
XX
PF 07-APR-2000; 2000WO-US09312.
XX
PR 09-APR-1999; 99US-0289198.
PR 28-OCT-1999; 99US-0429755.
PR 23-MAR-2000; 2000US-0534825.
XX
PA (CORI-) CORIXA CORP.
XX
PI Frudakis TN, Smith JM, Reed SG, Mitcham JL, Retter MW, Dillon DC;
XX
DR WPI; 2000-628403/60.
DR P-PSDB; AAB28628.
XX
PT An isolated polypeptide comprising an immunogenic portion of a breast
PT tumor protein used for inhibiting the development of cancer, especially

PT breast cancer, and monitoring cancer progression in a patient -
XX
PS Claim 4; Page 177; 187pp; English.
XX
CC The present sequence is given in a specification relating to compositions
CC and methods for the treatment and diagnosis of breast cancer. Nucleotide
CC sequences that are preferentially expressed in breast tumour tissue, and
CC the polypeptides encoded by such nucleotide sequences, are used in
CC compositions and vaccines to inhibit the development of cancer,
CC especially breast cancer. The progression of a cancer may be monitored by
CC carrying out detection of tumour-specific antigens at subsequent time
CC points and comparing the results from the different time points.
CC CD4+ and/or CD8- T-Cells isolated from the cancer patient may be treated
CC with tumour-specific polypeptides, polynucleotides encoding the
CC polypeptides or antigen presenting cells expressing the polypeptides. The
CC cells are then administered to the patient to inhibit development of
CC cancer.
XX
SQ Sequence 1155 BP; 346 A; 253 C; 297 G; 259 T; 0 other;

Query Match 100.0%; Score 1155; DB 21; Length 1155;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1155; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 atggtggttgaggttgattccatgccggctgcctcttctgtgaagaagccatttggtctc 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 atggtggttgaggttgattccatgccggctgcctcttctgtgaagaagccatttggtctc 60

Qy      61 aggagcaagatggcaagtggtgctgccgttgcttccccctgctgcagggagagcggcaag 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 aggagcaagatggcaagtggtgctgccgttgcttccccctgctgcagggagagcggcaag 120

Qy     121 agcaacgtgggcacttctggagaccacgacgactctgctatgaagacactcaggagcaag 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 agcaacgtgggcacttctggagaccacgacgactctgctatgaagacactcaggagcaag 180

Qy     181 atgggcaagtggtgccgccactgcttccccctgctgcaggggagtggcaagagcaacgtg 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 atgggcaagtggtgccgccactgcttccccctgctgcaggggagtggcaagagcaacgtg 240

Qy     241 ggcgcttctggaggaccacgacgactctgctatgaagacactcaggaacaagatgggcaag 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 ggcgcttctggaggaccacgacgactctgctatgaagacactcaggaacaagatgggcaag 300

Qy     301 tggtgctgccactgcttccccctgctgcaggggagcggcaagagcaaggtgggcgcttgg 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 tggtgctgccactgcttccccctgctgcaggggagcggcaagagcaaggtgggcgcttgg 360

Qy     361 ggagactacgatgacagtgccttcatggaggccaggtaccacgtccgtggagaagatctg 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 ggagactacgatgacagtgccttcatggaggccaggtaccacgtccgtggagaagatctg 420

Qy     421 gacaagctccacagagctgcctggtgtgggtaaagtccccagaaaggatctcatcgtcatg 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 gacaagctccacagagctgcctggtgtgggtaaagtccccagaaaggatctcatcgtcatg 480

Qy     481 ctcaggggacctgacgtgaacaagaaggacaagcaaaagaggactgctctacatctggcc 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 ctcaggggacctgacgtgaacaagaaggacaagcaaaagaggactgctctacatctggcc 540

Qy     541 tctgccaatgggaattcagaagtagtaaaactcctgctggacagacgatgtcaacttaat 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 tctgccaatgggaattcagaagtagtaaaactcctgctggacagacgatgtcaacttaat 600

Qy     601 gtccttgacaacaaaaagaaggacagctctgataaaggccgtacaatgccaggaagatgaa 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 gtccttgacaacaaaaagaaggacagctctgataaaggccgtacaatgccaggaagatgaa 660

Qy     661 tgtgcgttaatgttgctggaacatggcactgatccaaatattccagatgagtatggaaat 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 tgtgcgttaatgttgctggaacatggcactgatccaaatattccagatgagtatggaaat 720

Qy     721 accactctgcactacgctatctataatgaagataaattaatggccaaagcactgctctta 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 accactctgcactacgctatctataatgaagataaattaatggccaaagcactgctctta 780

Qy     781 tatggtgctgatatcgaatcaaaaaacaagcatggcctcacaccactgttacttggtgta 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 tatggtgctgatatcgaatcaaaaaacaagcatggcctcacaccactgttacttggtgta 840

Qy     841 catgagcaaaaacagcaagtcgtgaaatttttaatcaagaaaaaagcgaatttaaatgca 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 catgagcaaaaacagcaagtcgtgaaatttttaatcaagaaaaaagcgaatttaaatgca 900

Qy     901 ctggatagatatggaaggactgctctcatacttgctgtatgttgtggatcagcaagtata 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 ctggatagatatggaaggactgctctcatacttgctgtatgttgtggatcagcaagtata 960

Qy     961 gtcagccttctacttgagcaaaatattgatgtatcttctcaagatctatctggacagacg 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 gtcagccttctacttgagcaaaatattgatgtatcttctcaagatctatctggacagacg 1020

Qy    1021 gccagaagaagtatgctgtttctagtcatcatcatgtaatttgccagttacttctctgactac 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 gccagaagaagtatgctgtttctagtcatcatcatgtaatttgccagttacttctctgactac 1080

Qy    1081 aaagaaaaacagatgctaaaaatatcttctgaaaacagcaatccagaaaatgtctcaaga 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 aaagaaaaacagatgctaaaaatatcttctgaaaacagcaatccagaaaatgtctcaaga 1140

Qy    1141 accagaaataaataa 1155
          ||||||||||||||||
Db    1141 accagaaataaataa 1155
```

RESULT 2
AAA06598
ID AAA06598 standard; cDNA; 1155 BP.
XX
AC AAA06598;
XX
DT 13-JUN-2000 (first entry)
XX
DE Human immunogenic prostate tumour protein cDNA sequence SEQ ID NO:373.
XX
KW Human; prostate cancer; diagnosis, tumour, gene therapy, detection;
KW immunogenic; cytostatic; vaccine; ss
XX
OS Homo sapiens.
XX
PN WO200004149-A2.
XX
PD 27-JAN-2000.
XX
PF 14-JUL-1999; 99WO-US15838.
XX
PR 14-JUL-1998; 98US-0115453
PR 14-JUL-1998; 98US-0116134.
PR 23-SEP-1998; 98US-0159812.
PR 23-SEP-1998; 98US-0159822.
PR 15-JAN-1999; 99US-0232149.
PR 15-JAN-1999; 99US-0232880.
PR 09-APR-1999; 99US-0288946.
XX
PA (CORI-) CORIXA CORP.
XX
PI Dillon DC, Harlocker SL, Yuqiu J, Xu J, Mitcham JL,
XX
DR WPI; 2000-171268/15.
XX
PT New polypeptide useful for treating and diagnosing prostate cancer
PT comprises an immunogenic portion of prostate tumor protein -
XX
PS Claim 50; Page 222; 263pp; English.
XX

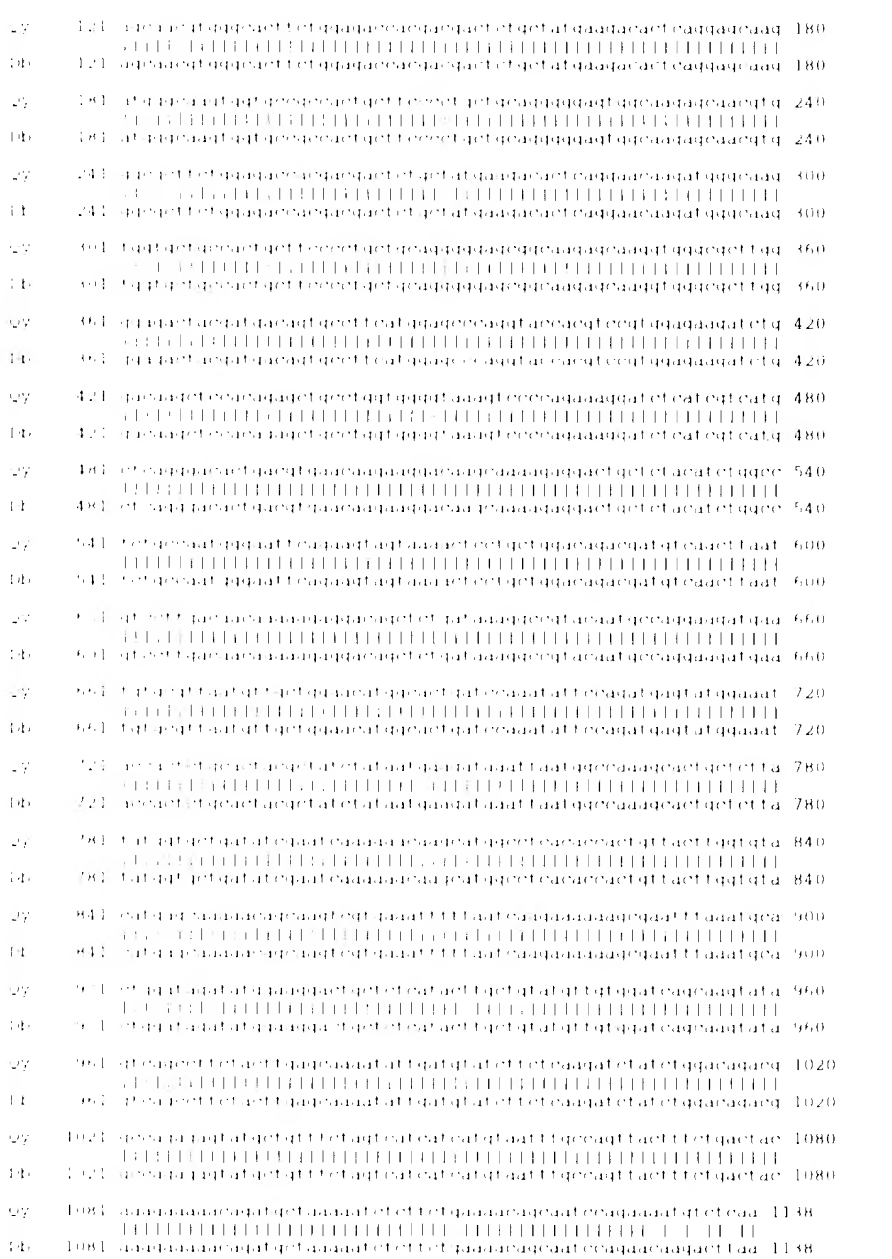
```
CC   The present invention describes isolated polypeptides, comprising an
CC   immunogenic portion of a prostate tumour protein (PTP). The polypeptides
CC   and polynucleotides encoding them have cytostatic activity and can be
CC   used in vaccines and in gene therapy. The polypeptides and
CC   polynucleotides encoding them, antigen presenting cells which express
CC   the polypeptides, antibodies against the polypeptides and vaccines
CC   comprising them can be used for inhibiting the development of prostate
CC   cancer in a patient. The polypeptides can be used to generate antibodies
CC   or anti idiotypic antibodies for passive immuno therapy. A portion of
CC   the polynucleotides encoding the polypeptides can be used as a probe or
CC   to modulate the expression of the polypeptides. AAA06241 to AAA06691 and
CC   AAY82... to AAY82... represent sequences used in the exemplification of
CC   the present invention.
XX
SQ   Sequence 1155 BP; 346 A; 253 C; 297 G; 259 T; 0 other;

  Query Match             100.0%;  Score 1155;  DB 21;  Length 1155;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1155;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 atggtggttgaggttgattccatgccggctgcctcttctgtgaagaagccatttggtctc 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 atggtggttgaggttgattccatgccggctgcctcttctgtgaagaagccatttggtctc 60

Qy         61 aggagcaagatgggcaagtggtgctgccgttgcttcccctgctgcaggggagagcggcaag 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 aggagcaagatgggcaagtggtgctgccgttgcttcccctgctgcaggggagagcggcaag 120

Qy        121 agcaacgtgggcacttctggagaccacgacgactctgctatgaagacactcaggagcaag 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 agcaacgtgggcacttctggagaccacgacgactctgctatgaagacactcaggagcaag 180

Qy        181 atggggaagtggtgccgccactgcttcccctgctgcagggggagtggcaagagcaacgtg 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 atggggaagtggtgccgccactgcttcccctgctgcagggggagtggcaagagcaacgtg 240

Qy        241 ggcacttctggagaccacgacgactctgctatgaagacactcaggaacaagatgggcaag 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ggcacttctggagaccacgacgactctgctatgaagacactcaggaacaagatgggcaag 300

Qy        301 tggtgctgccactgcttcccctgctgcaggggaagcggcaagagcaaggtgggcgcttgg 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 tggtgctgccactgcttcccctgctgcaggggaagcggcaagagcaaggtgggcgcttgg 360

Qy        361 ggagactacgatgacagtgccttcatggagcccaggtaccacgtccgtggagaagatctg 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ggagactacgatgacagtgccttcatggagcccaggtaccacgtccgtggagaagatctg 420

Qy        421 gacaagctccacagagctgcctggtggggtaaagtccccagaaaggatctcatcgtcatg 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 gacaagctccacagagctgcctggtggggtaaagtccccagaaaggatctcatcgtcatg 480

Qy        481 ctcagggacactgacgtgaacaagaaggacaagcaaaagaggactgctctacatctggcc 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ctcagggacactgacgtgaacaagaaggacaagcaaaagaggactgctctacatctggcc 540

Qy        541 tctgccaatgggaattcagaagtagtaaaactcctgctggacagacgatgtcaacttaat 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 tctgccaatgggaattcagaagtagtaaaactcctgctggacagacgatgtcaacttaat 600

Qy        601 gtccttgacaacaaaaagaggacagctctgataaaggccgtacaatgccaggaagatgaa 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 gtccttgacaacaaaaagaggacagctctgataaaggccgtacaatgccaggaagatgaa 660

Qy        661 tgtgcgttaatgttgctggaacatggcactgatccaaatattccagatgagtatggaaat 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 tgtgcgttaatgttgctggaacatggcactgatccaaatattccagatgagtatggaaat 720

Qy        721 accactctgcactacgctatctataatgaagataaattaatggccaaagcactgctctta 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 accactctgcactacgctatctataatgaagataaattaatggccaaagcactgctctta 780
```

```
Qy        781 tatggtgctgatatcgaatcaaaaaacaagcatggcctcacaccactgttacttggtgta 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 tatggtgctgatatcgaatcaaaaaacaagcatggcctcacaccactgttacttggtgta 840

Qy        841 catgagcaaaaacagcaagtcgtgaaatttttaatcaagaaaaaagcgaatttaaatgca 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 catgagcaaaaacagcaagtcgtgaaatttttaatcaagaaaaaagcgaatttaaatgca 900

Qy        901 ctggatagatatggaaggactgctctcatacttgctgtatgttgtggatcagcaagtata 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 ctggatagatatggaaggactgctctcatacttgctgtatgttgtggatcagcaagtata 960

Qy        961 gtcagccttctacttgagcaaaatattgatgtatcttctcaagatctatctggacagacg 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 gtcagccttctacttgagcaaaatattgatgtatcttctcaagatctatctggacagacg 1020

Qy       1021 gccagagagtatgctgtttctagtcatcatcatgtaatttgccagttacttctgactac 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 gccagagagtatgctgtttctagtcatcatcatgtaatttgccagttacttctgactac 1080

Qy       1081 aaagaaaaacagatgctaaaaatctcttctgaaaacagcaatccagaaaatgtctcaaga 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 aaagaaaaacagatgctaaaaatctcttctgaaaacagcaatccagaaaatgtctcaaga 1140

Qy       1141 accagaaataaataa 1155
              |||||||||||||||
Db       1141 accagaaataaataa 1155


RESULT 3
AAH93714
ID   AAH93714 standard; cDNA; 1155 BP.
XX
AC   AAH93714;
XX
DT   04-OCT-2001  (first entry)
XX
DE   Human prostate specific cDNA sequence B305D splice variant #8.
XX
KW   Human; prostate cancer; prostate specific; diagnosis; vaccine;
KW   cytostatic; gene therapy; metastasis; ss.
XX
OS   Homo sapiens.
XX
PN   WO200151633 A2.
XX
PD   19-JUL-2001.
XX
PF   16-JAN-2001; 2001WO-US01574.
XX
PR   14-JAN-2000; 2000US-0483672.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Xu J,  Dillon DC,  Mitcham JL,  Harlocker SL,  Jiang Y,  Reed SG;
PI   Kalos MD,  Fanger GR,  Day CH,  Retter MW,  Stolk JA,  Skeiky YAW;
PI   Wang A,  Meagher MJ;
XX
DR   WPI; 2001-425873/45.
XX
PT   New polynucleotide encoding a prostate specific protein, for
PT   diagnosing, monitoring and treating prostate cancer in a patient and
PT   for use in vaccines
XX
PS   Claim 1; Page 347; 543pp; English.
XX
CC   The present invention describes polynucleotide sequences (I) which encode
CC   prostate specific proteins (II). (I) and (II) have cytostatic activity,
CC   and can be used in vaccine production and gene therapy. (I), (II),
CC   antibodies to (II), fusion proteins comprising (II) and isolated
CC   T cells prepared using (I) or (II) are used treat cancer in a patient.
CC   (I) and the antibodies are also used in the detection of cancer in a
CC   patient. The cancer that is diagnosed or treated is particularly
```

```
CC   prostate cancer. (I) and (II) can be used in vaccines. The antibodies of
CC   (I) can be used for monitoring the progression of cancer in a patient.
CC   (I) and (II) can also be used to improve diagnostic and therapeutic
CC   methods for prostate cancer. They can indicate the level of metastasis
CC   as well as the prostate volume. AAH93357 to AAH93944 and AAM01115 to
CC   AAM01318 represent polynucleotide and amino acid sequences used in the
CC   exemplification of the present invention.
XX
SQ   Sequence 1155 BP; 346 A; 253 C; 297 G; 259 T; 0 other;

  Query Match             100.0%;  Score 1155;  DB 22;  Length 1155;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1155;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 atggtggttgaggttgattccatgccggctgcctcttctgtgaagaagccatttggtctc 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 atggtggttgaggttgattccatgccggctgcctcttctgtgaagaagccatttggtctc 60

Qy         61 aggagcaagatgggcaagtggtgctgccgttgcttcccctgctgcagggagagcggcaag 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 aggagcaagatgggcaagtggtgctgccgttgcttcccctgctgcagggagagcggcaag 120

Qy        121 agcaacgtgggcacttctggagaccacgacgactctgctatgaagacactcaggagcaag 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 agcaacgtgggcacttctggagaccacgacgactctgctatgaagacactcaggagcaag 180

Qy        181 atgggcaagtggtgccgccactgcttcccctgctgcagggggagtggcaagagcaacgtg 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 atgggcaagtggtgccgccactgcttcccctgctgcagggggagtggcaagagcaacgtg 240

Qy        241 ggcgcttctggagaccacgacgactctgctatgaagacactcaggaacaagatgggcaag 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ggcgcttctggagaccacgacgactctgctatgaagacactcaggaacaagatgggcaag 300

Qy        301 tggtgctgccactgcttcccctgctgcagggggagcggcaagagcaaggtgggcgcttgg 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 tggtgctgccactgcttcccctgctgcagggggagcggcaagagcaaggtgggcgcttgg 360

Qy        361 ggagactacgatgacagtgccttcatggagcccaggtaccacgtccgtggagaagatctg 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ggagactacgatgacagtgccttcatggagcccaggtaccacgtccgtggagaagatctg 420

Qy        421 gacaagctccacagagctgcctggtggggtaaagtccccagaaaggatctcatcgtcatg 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 gacaagctccacagagctgcctggtggggtaaagtccccagaaaggatctcatcgtcatg 480

Qy        481 ctcagggacactgacgtgaacaagaaggacaagcaaaagaggactgctctacatctggcc 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ctcagggacactgacgtgaacaagaaggacaagcaaaagaggactgctctacatctggcc 540

Qy        541 tctgccaatgggaattcagaagtagtaaaactcctgctggacagacgatgtcaacttaat 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 tctgccaatgggaattcagaagtagtaaaactcctgctggacagacgatgtcaacttaat 600

Qy        601 gtccttgacaacaaaaagaggacagctctgataaaggccgtacaatgccaggaagatgaa 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 gtccttgacaacaaaaagaggacagctctgataaaggccgtacaatgccaggaagatgaa 660

Qy        661 tgtgcgttaatgttgctggaacatggcactgatccaaatattccagatgagtatggaaat 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 tgtgcgttaatgttgctggaacatggcactgatccaaatattccagatgagtatggaaat 720

Qy        721 accactctgcactacgctatctataatgaggataaattaatggccaaagcactgctctta 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 accactctgcactacgctatctataatgaggataaattaatggccaaagcactgctctta 780

Qy        781 tatggtgctgatatcgaatcaaaaaacaagcatggcctcacaccactgttacttggtgta 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 tatggtgctgatatcgaatcaaaaaacaagcatggcctcacaccactgttacttggtgta 840

Qy        841 catgagcaaaaacagcaagtcgtgaaatttttaatcaagaaaaaagcgaatttaaatgca 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 catgagcaaaaacagcaagtcgtgaaatttttaatcaagaaaaaagcgaatttaaatgca 900

Qy        901 ctggatagatatggaaggactgctctcatacttgctgtatgttgtggatcagcaagtata 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 ctggatagatatggaaggactgctctcatacttgctgtatgttgtggatcagcaagtata 960

Qy        961 gtcagccttctacttgagcaaaatattgatgtatcttctcaagatctatctggacagacg 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 gtcagccttctacttgagcaaaatattgatgtatcttctcaagatctatctggacagacg 1020

Qy       1021 gccagagagtatgctgtttctagtcatcatcatgtaatttgccagttactttctgactac 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 gccagagagtatgctgtttctagtcatcatcatgtaatttgccagttactttctgactac 1080

Qy       1081 aaagaaaaacagatgctaaaaatctcttctgaaaacagcaatccagaaaatgtctcaaga 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 aaagaaaaacagatgctaaaaatctcttctgaaaacagcaatccagaaaatgtctcaaga 1140

Qy       1141 accagaaataaataa 1155
              |||||||||||||||
Db       1141 accagaaataaataa 1155


RESULT 4
AAH85028
ID   AAH85028 standard; cDNA; 1155 BP.
XX
AC   AAH85028;
XX
DT   25-SEP-2001  (first entry)
XX
DE   Human prostate-specific cDNA sequence B305D splice variant #8.
XX
KW   Human; prostate cancer; therapy; diagnosis; cat eye syndrome;
KW   chromosome 22q11.2; prostate-specific protein; chromosome 1;
KW   prostate specific antigen, PSA; ss.
XX
OS   Homo sapiens.
XX
PN   WO200134802-A2.
XX
PD   17-MAY-2001.
XX
PF   09-NOV-2000; 2000WO-US30904.
XX
PR   12-NOV-1999;   99US-0439313.
PR   18-NOV-1999;   99US-0443686.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Xu J,  Dillon DC,  Mitcham JL,  Harlocker SL,  Jiang Y,  Reed SG;
PI   Kalos MD,  Fetter MW,  Stolk JA,  Day CH,  Skeiky YAW,  Wang A;
XX
DR   WPI; 2001-308785/32.
XX
PT   Isolated polypeptide comprising at least an immunogenic portion of a
PT   prostate-specific protein, useful in the diagnosis and therapy of
PT   prostate cancer -
XX
PS   Claim 31; Page 246-247; 325pp; English.
XX
CC   The present invention describes an isolated polypeptide (PI) comprising
CC   at least an immunogenic portion of a prostate-specific protein, or its
CC   variant. Also described are polynucleotides (NI) encoding (PI). (PI) and
CC   (NI) have cytostatic activity and can be used in vaccine production.
CC   The polypeptides, nucleic acids and antibodies from the present
CC   invention are useful in the diagnosis and therapy of prostate cancer.
CC   Prostate specific genes P704P, P712P, P774P, P775P and B305D are located
CC   in a genomic region on chromosome 22q11.2 known as the Cat Eye Syndrome
CC   region. Prostate specific antigen (PSA) P501S was located on
CC   chromosome 1. AAH84671 to AAH85143 and AAG99000 to AAG99077 represent
CC   polynucleotide and polypeptide sequences used in the exemplification
CC   of the present invention.
```

```
XX
SQ   Sequence 1155 BP; 346 A; 253 C; 297 G; 259 T; 0 other;

  Query Match             100.0%;  Score 1155;  DB 22;  Length 1155;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1155;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 atggtggttgaggttgattccatgcccgctgcctctttctgtgaagaagccatttggtctc 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 atggtggttgaggttgattccatgcccgctgcctctttctgtgaagaagccatttggtctc 60

Qy         61 aggaagcaagatgggcaagtggtgctgccgtttgcttcccctgctgcaggggagagcggcaag 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 aggaagcaagatgggcaagtggtgctgccgtttgcttcccctgctgcaggggagagcggcaag 120

Qy        121 agcaacgtgggcacttctggagaccacgacgactctgctatgaagacactcaggagcaag 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 agcaacgtgggcacttctggagaccacgacgactctgctatgaagacactcaggagcaag 180

Qy        181 atgggcaagtggtgccgccactgcttcccctgctgcaggggagtggcaagagcaacgtg 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 atgggcaagtggtgccgccactgcttcccctgctgcaggggagtggcaagagcaacgtg 240

Qy        241 ggcgcttctggagaccacgacgactctgctatgaagacactcaggaacaagatgggcaag 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ggcgcttctggagaccacgacgactctgctatgaagacactcaggaacaagatgggcaag 300

Qy        301 tggtgctgccactgcttcccctgctgcagggggagcggcaagagcaagtgggcgcttgg 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 tggtgctgccactgcttcccctgctgcagggggagcggcaagagcaagtgggcgcttgg 360

Qy        361 gggagctacgatgacagtgccttcatggagcccaggtaccacgtccgtggagaagatctg 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 gggagctacgatgacagtgccttcatggagcccaggtaccacgtccgtggagaagatctg 420

Qy        421 gacaagctccacagagctgcctggtgggtaaagtccccagaaaggatctcatcgtcatg 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 gacaagctccacagagctgcctggtgggtaaagtccccagaaaggatctcatcgtcatg 480

Qy        481 ctcagggacactgacgtgaacaagaaggacaagcaaaagaggactgctctacatctggcc 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ctcagggacactgacgtgaacaagaaggacaagcaaaagaggactgctctacatctggcc 540

Qy        541 tctgccaatgggaattcagaagtagtaaaactcctgctggacagacgatgtcaacttaat 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 tctgccaatgggaattcagaagtagtaaaactcctgctggacagacgatgtcaacttaat 600

Qy        601 gtccttgacaacaaaaagaggacagctctgataaaggccgtacaatgccaggaagatgaa 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 gtccttgacaacaaaaagaggacagctctgataaaggccgtacaatgccaggaagatgaa 660

Qy        661 tgtgcgttaatgttgctggaacatggcactgatccaaatattccagatgagtatggaaat 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 tgtgcgttaatgttgctggaacatggcactgatccaaatattccagatgagtatggaaat 720

Qy        721 accactctgcactacgctatctataatgaagataaattaatggccaaagcactgctctta 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 accactctgcactacgctatctataatgaagataaattaatggccaaagcactgctctta 780

Qy        781 tatggtgctgatatcgaatcaaaaaacaagcatggcctcacaccactgttacttggtgta 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 tatggtgctgatatcgaatcaaaaaacaagcatggcctcacaccactgttacttggtgta 840

Qy        841 catgagcaaaaacagcaagtcgtgaaatttttaatcaagaaaaaagcgaatttaaatgca 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 catgagcaaaaacagcaagtcgtgaaatttttaatcaagaaaaaagcgaatttaaatgca 900

Qy        901 ctggatagatatggaaggactgctctcatacttgctgtatgttgtggatcagcaagtata 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 ctggatagatatggaaggactgctctcatacttgctgtatgttgtggatcagcaagtata 960

Qy        961 gtcagccttctacttgagcaaaatattgatgtatcttctcaagatctatctggacagacg 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 gtcagccttctacttgagcaaaatattgatgtatcttctcaagatctatctggacagacg 1020

Qy       1021 gccagagaatatgctgtttctagtcatcatcatgtaatttgccagttacttctgactac 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 gccagagaatatgctgtttctagtcatcatcatgtaatttgccagttacttctgactac 1080

Qy       1081 aaagaaaaacagatgctaaaaatctcttctgaaaacagcaatccagaaaatgtctcaaga 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 aaagaaaaacagatgctaaaaatctcttctgaaaacagcaatccagaaaatgtctcaaga 1140

Qy       1141 accagaaataaataa 1155
              |||||||||||||||
Db       1141 accagaaataaataa 1155


RESULT    5
AAH02779
ID   AAH02779 standard; cDNA; 1155 BP.
XX
AC   AAH02779;
XX
DT   14-JUN-2001  (first entry)
XX
DE   Prostate tumour antigen determined cDNA splice variant of B305D #e
XX
KW   Human; prostate tumour antigen; prostate tumour; therapy; diagnosis;
KW   prostate cancer; immunogenic; cytostatic; vaccine; ss.
XX
OS   Homo sapiens.
XX
PN   WO200125272-A2.
XX
PD   12-APR-2001.
XX
PF   04-OCT-2000; 2000WO-US27464.
XX
PR   04-OCT-1999;   99US-0157455.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Xu J,  Skeiky YAW,  Reed SG,  Cheever MA;
XX
DR   WPI; 2001-245062/25.
DR   P-PSDB; AAB74815.
XX
PT   Prostate specific protein and its encoding polynucleotide, useful for
PT   the treatment and diagnosis of prostate cancer
XX
PS   Claim 50; Page 231-232; 276pp; English.
XX
CC   The present invention describes an isolated polypeptide (I) comprising
CC   at least an immunogenic portion of a prostate tumour antigen protein or
CC   its variant. (I) have cytostatic activity and can be used in vaccine
CC   production. (I), prostate tumour antigen polynucleotides, an antigen
CC   presenting cell (APC e.g. a dendritic cell) that expresses (I), and a
CC   pharmaceutical composition containing (I) are useful for inhibiting the
CC   development of cancer in a patient. Antibodies specific for prostate
CC   specific proteins and oligonucleotides that hybridise to a
CC   polynucleotide that encodes a prostate specific protein are useful
CC   for detecting the presence or absence of a cancer or monitoring the
CC   progression the progression of a cancer, especially prostate cancer.
CC   AAH02422 to AAH2872, AAB74798 to AAB74821 and AAB74966 are sequences
CC   used in the exemplification of the present invention.
XX
SQ   Sequence 1155 BP; 346 A; 253 C; 297 G; 259 T; 0 other;

  Query Match             100.0%;  Score 1155;  DB 22;  Length 1155;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1155;  Conservative    0;  Mismatches    0;  Indels    0;  Gap    0;

Qy          1 atggtggttgaggttgattccatgcccgctgcctctttctgtgaagaagccatttggtctc
```

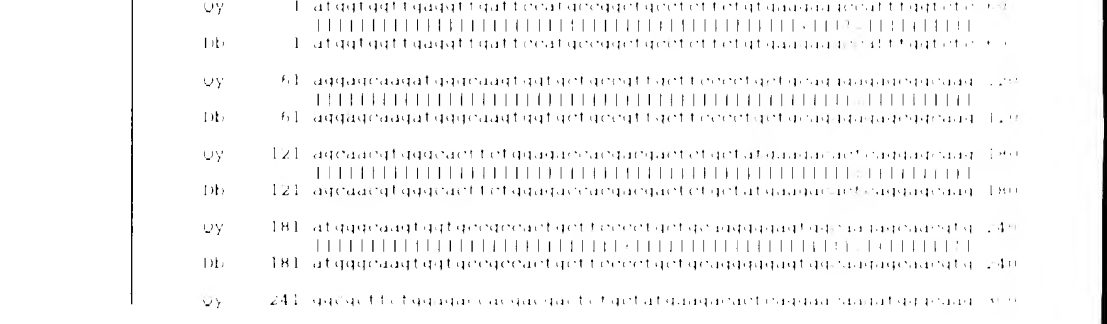
```
Db      1 atggtggttgaggttgattccatgccggctgcctcttctgtgaagaagccatttggtctc 60

Qy     61 aggagcaagatgggcaagtggtgctgccgttgcttcccctgctgcagggagaggcggcaag 120
Db     61 aggagcaagatgggcaagtggtgctgccgttgcttcccctgctgcagggagaggcggcaag 120

Qy    121 agcaacgtgggcacttctggagaccacgacgacttctgctatgaagacactcaggaggcaag 180
Db    121 agcaacgtgggcacttctggagaccacgacgacttctgctatgaagacactcaggaggcaag 180

Qy    181 atgggcaagtggtaccgccgccactgcttcccctgctgcagggagggagtggcaagagcaacg 240
Db    181 atgggcaagtggtaccgccgccactgcttcccctgctgcagggagggagtggcaagagcaacg 240

Qy    241 ggcgcttctggaggagcagcggacgacttctgctatgaagacactccaggaacaagatgggcaag 300
Db    241 ggcgcttctggaggagcagcggacgacttctgctatgaagacactccaggaacaagatgggcaag 300

Qy    301 tggtgctgccactgcttcccctgctgcagggggagcgggcaagagcaaggtgggcgcttgg 360
Db    301 tggtgctgccactgcttcccctgctgcagggggagcgggcaagagcaaggtgggcgcttgg 360

Qy    361 ggaggacttacgatgacagtgccttcatggaggccaggtaccacgtccgtggagaagatctg 420
Db    361 ggaggacttacgatgacagtgccttcatggaggccaggtaccacgtccgtggagaagatctg 420

Qy    421 gacaaagctccacaagagctgcctggtgggtaaagtcccagaaagatctcatcgtcatg 480
Db    421 gacaaagctccacaagagctgcctggtgggtaaagtcccagaaagatctcatcgtcatg 480

Qy    481 ctcagggacactgacgtgaacaagaagcgacaagcaaaagaggactgctctacatctggcc 540
Db    481 ctcagggacactgacgtgaacaagaagcgacaagcaaaagaggactgctctacatctggcc 540

Qy    541 tctgccaatgaggaattcagaagtagtaaaactcctgctgcacagacgatgtcaacttaat 600
Db    541 tctgccaatgaggaattcagaagtagtaaaactcctgctgcacagacgatgtcaacttaat 600

Qy    601 gtccttgacaacaaaaagaggacagctctgataaagccgtacaatgccagggaagatgaa 660
Db    601 gtccttgacaacaaaaagaggacagctctgataaagccgtacaatgccagggaagatgaa 660

Qy    661 tgtgcgttaatgttgctgtgaaacatggcactgatccaaatattccagatgagtatggaaat 720
Db    661 tgtgcgttaatgttgctgtgaaacatggcactgatccaaatattccagatgagtatggaaat 720

Qy    721 accactctgcactacgctatctataatgaagataaattaatgggcaaagcactgctctta 780
Db    721 accactctgcactacgctatctataatgaagataaattaatgggcaaagcactgctctta 780

Qy    781 tatggtgctgatatcgaatcaaaaaacaagcatggctcacaccactgttacttggtgta 840
Db    781 tatggtgctgatatcgaatcaaaaaacaagcatggctcacaccactgttacttggtgta 840

Qy    841 catgagcaaaaacagcaagtcgtgaaatttttaatcaagaaaaaagcgaatttaaatgca 900
Db    841 catgagcaaaaacagcaagtcgtgaaatttttaatcaagaaaaaagcgaatttaaatgca 900

Qy    901 ctggatagatatgaaaggactgctctcatacttgctgtatgttgtggatcagcaagtata 960
Db    901 ctggatagatatgaaaggactgctctcatacttgctgtatgttgtggatcagcaagtata 960

Qy    961 gtcagccttctacttgagcaaaatattgatgtatcttctcaagatctatctggacagacg 1020
Db    961 gtcagccttctacttgagcaaaatattgatgtatcttctcaagatctatctggacagacg 1020

Qy   1021 gccagaagtatgctgtttctagtcatcatcatgtaatttgccagttacttttctgactac 1080
Db   1021 gccagaagtatgctgtttctagtcatcatcatgtaatttgccagttacttttctgactac 1080

Qy   1081 aaagaaaaacagatgctaaaaatctcttctgaaaacagcaatccagaaaatgtctcaaga 1140
Db   1081 aaagaaaaacagatgctaaaaatctcttctgaaaacagcaatccagaaaatgtctcaaga 1140

Qy   1141 accagaaataaataa 1155
Db   1141 accagaaataaataa 1155


RESULT   6
AAC81012
ID   AAC81012 standard; cDNA; 2000 BP.
XX
AC   AAC81012;
XX
DT   13-FEB-2001  (first entry)
XX
DE   Human B11Aq1 antigen splice isoform B11C-8 cDNA.
XX
KW   Human; breast tumour specific antigen, cytostatic, vaccine;
KW   breast cancer; B18Aq1; B11Aq1; B15Aq1; ss.
XX
OS   Homo sapiens.
XX
PN   WO200061753-A2.
XX
PD   19-OCT-2000.
XX
PF   07-APR-2000; 2000WO-US09312.
XX
PR   09-APR-1999;   99US-0289198.
PR   28-OCT-1999;   99US-0429755.
PR   23-MAR-2000; 2000US-0534825.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Frudakis TN,  Smith JM,  Reed SG,  Misher LE,  Retter MW,  Dillon DC;
XX
DR   WPI; 2000-628403/60
DR   P-PSDB; AAB28629.
XX
PT   An isolated polypeptide comprising an immunogenic portion of a breast
PT   tumor protein used for inhibiting the development of cancer, especially
PT   breast cancer, and monitoring cancer progression in a patient
XX
PS   Claim 4; Page 177-178; 187pp; English.
XX
CC   The present sequence is given in a specification relating to compositions
CC   and methods for the treatment and diagnosis of breast cancer. Nucleotide
CC   sequences that are preferentially expressed in breast tumour tissue, and
CC   the polypeptides encoded by such nucleotide sequences, are used in
CC   compositions and vaccines to inhibit the development of cancer,
CC   especially breast cancer. The progression of a cancer may be monitored by
CC   carrying out detection of tumour-specific antigens at subsequent time
CC   points and comparing the results from the different time points.
CC   CD4+ and/or CD8+ T-Cells isolated from the cancer patient may be treated
CC   with tumour-specific polypeptides, polynucleotides encoding the
CC   polypeptides or antigen presenting cells expressing the polypeptides. The
CC   cells are then administered to the patient to inhibit development of
CC   cancer.
XX
SQ   Sequence 2000 BP; 698 A; 388 C; 489 G; 425 T; 0 other;


  Query Match             98.0%;  Score 1131.6;  DB 21;  Length 2000;
  Best Local Similarity   99.6%;  Pred. No. 0;
  Matches 1134;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 atggtggttgaggttgattccatgccggctgcctcttctgtgaagaagccatttggtctc 60
Db       1 atggtggttgaggttgattccatgccggctgcctcttctgtgaagaagccatttggtctc 60

Qy      61 aggagcaagatgggcaagtggtgctgccgttgcttcccctgctgcagggagaggcggcaag 120
Db      61 aggagcaagatgggcaagtggtgctgccgttgcttcccctgctgcagggagaggcggcaag 120
```



RESULT 7
AAH93715
ID AAH93715 standard; cDNA; 2000 BP.

XX
AC AAH93715;
XX
DT 04-OCT-2001 (first entry)
XX
DE Human prostate-specific cDNA sequence B405b splice variant #6.
XX
KW Human; prostate cancer; prostate specific; diagnosis; vaccine;
KW cytostatic; gene therapy; metastasis; ss.
XX
OS Homo sapiens.
XX
PN WO200151633-A2.
XX
PD 19-JUL-2001.
XX
PF 16-JAN-2001; 2001WO-US01574.
XX
PR 14-JAN-2000; 2000US-0483672.
XX
PA (CORI-) CORIXA CORP.
XX
PI Xu J, Dillon DC, Mitcham JL, Harlocker SL, Duan Y, Reed SG,
PI Kalos MD, Parker GK, Toy CB, Reiter RW, Stolk JA, Oberoy YAW
PI Wang A, Meagher MJ;
XX
DR WPI; 2001-425873/45.
XX
PT New polynucleotide encoding a prostate specific protein, for
PT diagnosing, monitoring and treating prostate cancer in a patient and
PT for use in vaccines -
XX
PS Claim 1; Page 347-348; 54pp; English.
XX
CC The present invention describes polynucleotide sequences (I) which encode
CC prostate specific proteins (II). (I) and (II) have cytostatic activity,
CC and can be used in vaccine production and gene therapy. (I), (II)
CC antibodies to (II), fusion proteins comprising (II), and isolated
CC T cells prepared using (I) or (II) are used treat cancer in a patient
CC (I) and the antibodies are also used in the detection of cancer in a
CC patient. The cancer that is diagnosed or treated is particularly
CC prostate cancer. (I) and (II) can be used in vaccines. The antibodies
CC (I) can be used for monitoring the progression of cancer in a patient
CC (I) and (II) can also be used to improve diagnostic and therapeutic
CC methods for prostate cancer. They can indicate the level of metastasis
CC as well as the prostate volume. AAH93657 to AAH93744 and AAM01317-1-
CC AAM01318 represent polynucleotide and amino acid sequences used in the
CC exemplification of the present invention.
XX
SQ Sequence 2000 BP; 698 A; 488 C; 489 G; 325 T; 0 other;

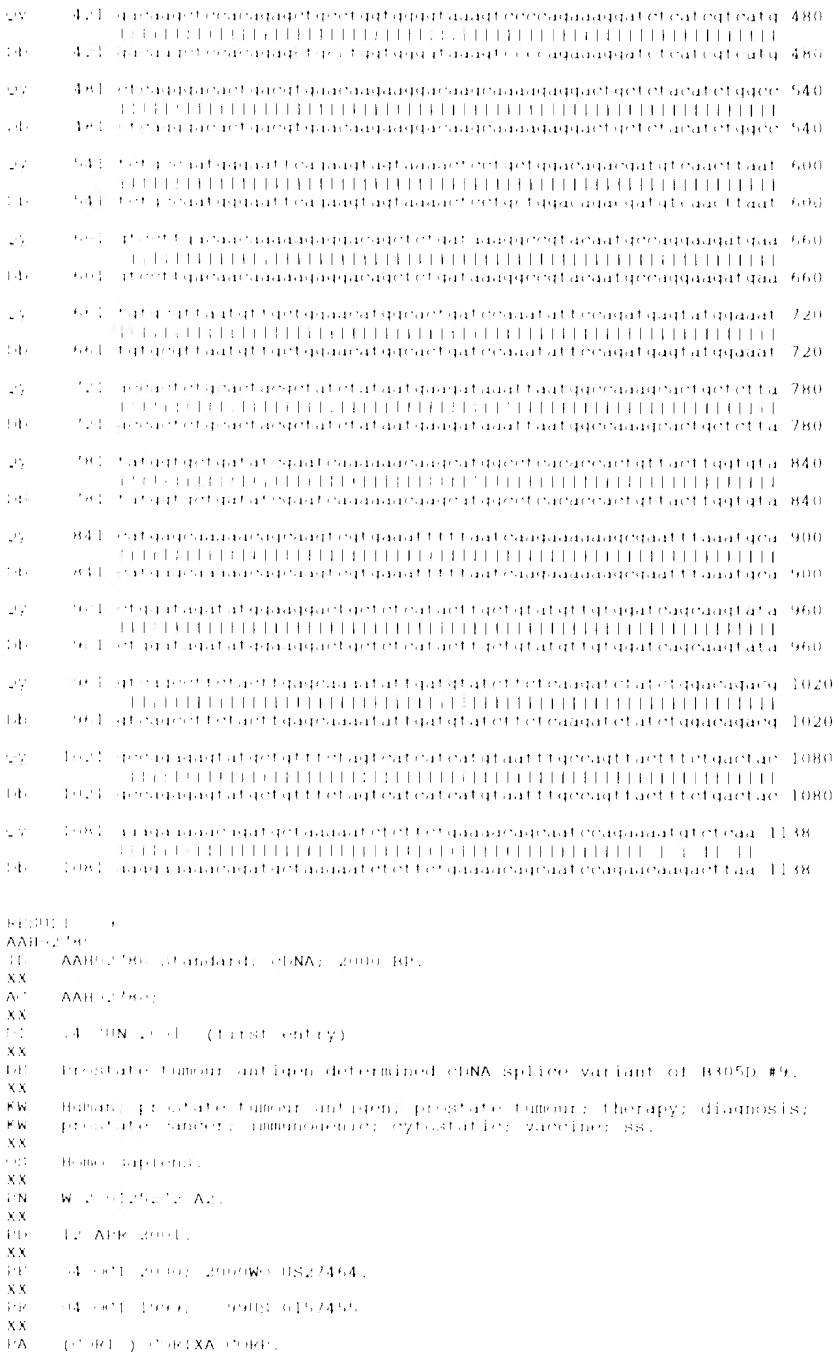
Query Match 98.0%; Score 1131.6; DB 22; Length 2000;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1134; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

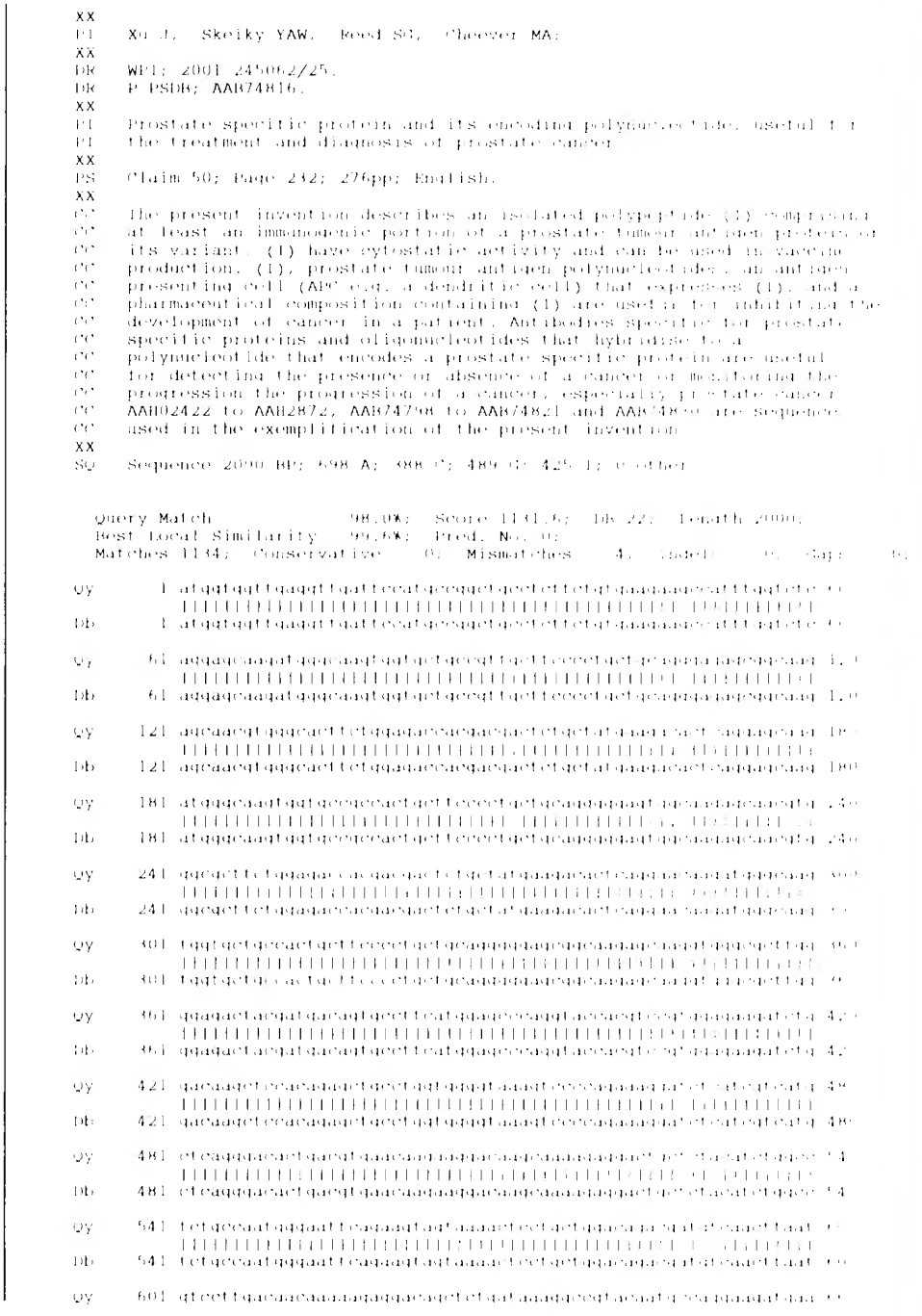


```
Db    241 qgcgcttctgggagaccacgacgactctgctatgaagacactcaggaacaagatgggcaag 300

Qy    301 tggtgctgccactgcttcccctgctgcagggggagcgggcaggagcaagtgggcgcttgg 360
Db    301 tggtgctgccactgcttcccctgctgcagggggagcgggcaggagcaagtgggcgcttgg 360

Qy    361 ggagactacgatgacagtgccttcatggagcccaggtaccacgtccgtgggagaagatctg 420
Db    361 ggagactacgatgacagtgccttcatggagcccaggtaccacgtccgtgggagaagatctg 420

Qy    421 gacaagctccacacagagctgcctggtgggatgaagtcccagaaaggatctcatcgtcatg 480
Db    421 gacaagctccacacagagctgcctggtgggatgaagtcccagaaaggatctcatcgtcatg 480

Qy    481 ctcaggacactgacgtgaacaaggacgacgcaaaaaggaggactgtctcacatctgccc 540
Db    481 ctcaggacactgacgtgaacaaggacgacgcaaaaaggaggactgtctcacatctgccc 540

Qy    541 tctgccaatgggaattcagaagtagtaaaactcctgctgcagacagacgatgtcaacttaat 600
Db    541 tctgccaatgggaattcagaagtagtaaaactcctgctgcagacagacgatgtcaacttaat 600

Qy    601 gtccttgacaacaaaaagaggacagctctgataaaggccgtacaatgccaggaagatgaa 660
Db    601 gtccttgacaacaaaaagaggacagctctgataaaggccgtacaatgccaggaagatgaa 660

Qy    661 tgtgcgttaatgttgctggaacatggcactgatccaaatattccagatgagtatggaaat 720
Db    661 tgtgcgttaatgttgctggaacatggcactgatccaaatattccagatgagtatggaaat 720

Qy    721 accactctgcactacgctatctataatgaagatgacattagtggccaaacgactgctctta 780
Db    721 accactctgcactacgctatctataatgaagatgacattagtggccaaacgactgctctta 780

Qy    781 tatgtgctgatatcgaatcaaaaaacaaacatggcctcacaccactgttacttggtgta 840
Db    781 tatgtgctgatatcgaatcaaaaaacaaacatggcctcacaccactgttacttggtgta 840

Qy    841 catgagcaaaaacagcaagtcgtgaaatttttaatcaagaaaaaagcgaatttaaatgca 900
Db    841 catgagcaaaaacagcaagtcgtgaaatttttaatcaagaaaaaagcgaatttaaatgca 900

Qy    901 ctggatagatatgagaaggactgctctcatacttgctgtatgttgtggatcagcaagtata 960
Db    901 ctggatagatatgagaaggactgctctcatacttgctgtatgttgtggatcagcaagtata 960

Qy    961 gtcagccttctacttgagcaaaatattgatgtatctttctcaagatctatctggacagacg 1020
Db    961 gtcagccttctacttgagcaaaatattgatgtatctttctcaagatctatctggacagacg 1020

Qy   1021 gccagagagtatgctgtttctagtcatcatcatgtaatttgccagttacttttctgactac 1080
Db   1021 gccagagagtatgctgtttctagtcatcatcatgtaatttgccagttacttttctgactac 1080

Qy   1081 aaagaaaaacagatactaaaaatctcttctgaaacagcaatccagaaaaatgtctcaa 1138
Db   1081 aaagaaaaacagatgctaaaaatctcttctgaaacagcaatccagaaacaagacttaa 1138
```

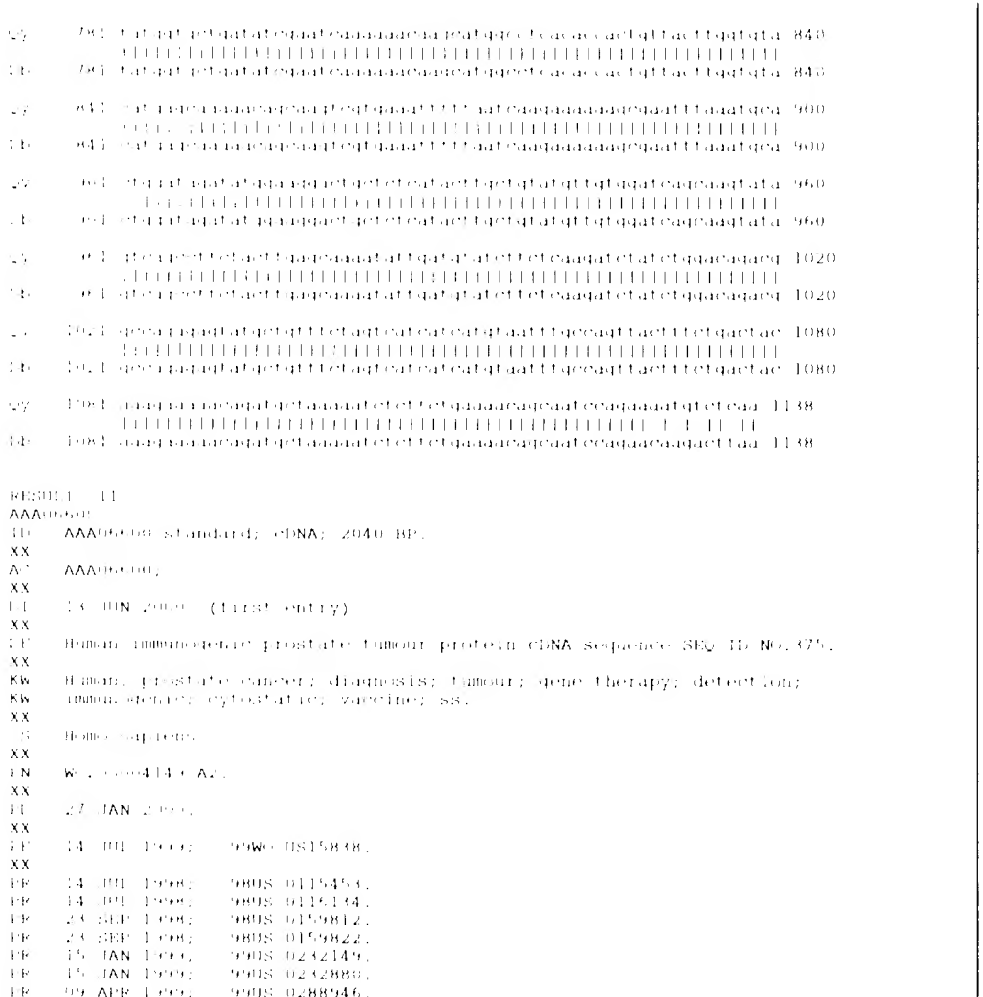
RESULT 8
AAH85029
ID AAH85029 standard; cDNA; 2000 BP.
XX
AC AAH85029;
XX
DT 25-SEP-2001 (first entry)
XX
DE Human prostate-specific cDNA sequence B305D splice variant #9.
XX
KW Human; prostate cancer; therapy; diagnosis; cat eye syndrome;
KW chromosome 22q11.2; prostate specific protein, chromosome 1;
KW prostate specific antigen; PSA; ss.
XX
OS Homo sapiens.
XX
PN WO200134802 A2.
XX
PD 17-MAY-2001.
XX
PF 09-NOV-2000; 2000WO-US30904.
XX
PR 12-NOV-1999; 99US-0439313.
PR 18-NOV-1999; 99US-0443686.
XX
PA (CORI-) CORIXA CORP.
XX
PI Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG;
PI Kales MP, Retter MW, Stolk JA, Day CH, Skolky YAW, Wang A;
XX
DR WPI; 2001-308785/32.
XX
PT Isolated polypeptide comprising at least an immunogenic portion of a
PT prostate specific protein, useful in the diagnosis and therapy of
PT prostate cancer -
XX
PS Claim 31; Page 247-248; 325pp; English.
XX
CC The present invention describes an isolated polypeptide (P1) comprising
CC at least an immunogenic portion of a prostate-specific protein, or its
CC variant. Also described are polynucleotides (N1) encoding (P1), (P1) and
CC (N1) have cytostatic activity and can be used in vaccine production.
CC The polypeptides, nucleic acids and antibodies from the present
CC invention are useful in the diagnosis and therapy of prostate cancer.
CC Prostate specific genes P704P, P712P, P774P, P775P and P509S are located
CC in a genetic region known as the Cat Eye Syndrome
CC region. Prostate specific antigen (PSA) P501S was located on
CC chromosome 1. AAH84671 to AAH85143 and AAG99000 to AAG99077 represent
CC polynucleotide and polypeptide sequences used in the exemplification
CC of the present invention.
XX
SQ Sequence 2000 BP; 698 A; 388 C; 489 G; 425 T; 0 other;

```
  Query Match          98.0%;  Score 1121.6;  DB 22;  Length 2000;
  Best Local Similarity  99.6%;  Pred. No. 0;
  Matches 1134;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 atggtggttgaggttgattccatgccggctgccttctgtgaagaagccatttggtctc 60
Db       1 atggtggttgaggttgattccatgccggctgccttctgtgaagaagccatttggtctc 60

Qy      61 aggagcaagatgggcaaqtggtgctgccgttgcttcccctgctgcagggagaggcgqcaag 120
Db      61 aggagcaagatgggcaagtggtgctgccgttgcttcccctgctgcagggagaggcggcaag 120

Qy     121 agcagctgggccacttctgggagaccacgacgactctgctatgaagacactcaggaacaag 180
Db     121 agcagctgggccacttctgggagaccacgacgactctgctatgaagacactcaggaacaag 180

Qy     181 atgggcaagtggtgccgccactgcttcccctgctgcagggggaggtggcaagagcaacgtg 240
Db     181 atgggcaagtggtgccgccactgcttcccctgctgcagggggaggtggcaagagcaacgtg 240

Qy     241 ggcgcttctgggagaccacgacgactctgctatgaagacactcaggaacaagatgggcaag 300
Db     241 ggcgcttctgggagaccacgacgactctgctatgaagacactcaggaacaagatgggcaag 300

Qy     301 tggtgctgccactgcttcccctgctgcagggggaggcggcaagagcaaggtggtgctgctt 360
Db     301 tggtgctgccactgcttcccctgctgcagggggaggcggcaagagcaaggtggtgctgctt 360

Qy     361 ggagactacgatgacagtgccttcatggagcccaggtaccacgtccgtgggagaagatctg 420
Db     361 ggagactacgatgacagtgccttcatggagcccaggtaccacgtccgtgggagaagatctg 420
```



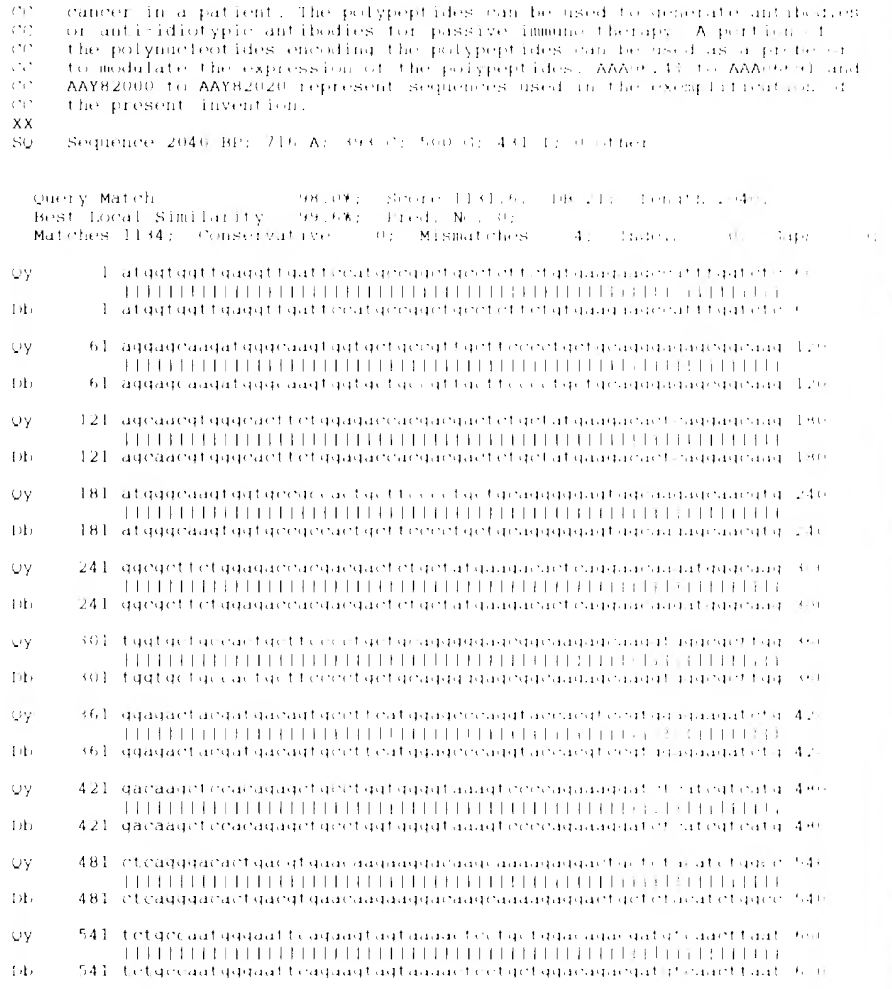


```
Db    601 gtccttgacaacaaaaagagggacagctctgataaaggccgtacaatgccagggaagatgaa 660

Qy    661 tgtgcgttaatgttgctggaacatggcaattgatctaaaatattccagatgagtatggaaat 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 tgtgcgttaatgttgctggaacatggcaattgatctaaaatattccagatgagtatggaaat 720

Qy    721 accactctgcactacgctatctataatgaaggataaattaattgtgccaaagcaattgctctta 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 accactctgcactacgctatctataatgaaggataaattaattgtgccaaagcaattgctctta 780

Qy    781 tatggtgctgatatcgaatcaaaaaacaagcatgagcctcacaccactgttacttggtgtta 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 tatggtgctgatatcgaatcaaaaaacaagcatgagcctcacaccactgttacttggtgtta 840

Qy    841 catgaagcaaaaaacagcaagtcgtgaaatttttaatcaagaaaaaagcgaatttaaaatgca 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 catgaagcaaaaaacagcaagtcgtgaaatttttaatcaagaaaaaagcgaatttaaaatgca 900

Qy    901 ctggatagatatgtgaaggactgctctcatacttgctgtatgttgtggatcagcaagtata 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 ctggatagatatgtgaaggactgctctcatacttgctgtatgttgtggatcagcaagtata 960

Qy    961 gtcagccttctacttgagcaaaatattgatgtatcttctcaagatctatctggacagacg 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 gtcagccttctacttgagcaaaatattgatgtatcttctcaagatctatctggacagacg 1020

Qy   1021 gccagaagagtatgctgtttctagtcatcatcatgtaatttgccagttacttttctgactac 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 gccagaagagtatgctgtttctagtcatcatcatgtaatttgccagttacttttctgactac 1080

Qy   1081 aaaaaaaaacagatgctaaaaatctcttctgaaaacagcaatccagaaaatgtctcaa 1138
          |||||||||||||||||||||||||||||||||||||||||||||| | | || ||
Db   1081 aaaaaaaaacagatgctaaaaatctcttctgaaaacagcaatccagaacaagacttaa 1138


RESULT 10
AAC81013
ID   AAC81013 standard; cDNA; 2040 BP.
XX
AC   AAC81013;
XX
DT   14-FEB-2001 (first entry)
XX
DE   Human B11Aq1 antigen splice isoform B11C-9-16 cDNA.
XX
KW   Human; breast tumour-specific antigen; cytostatic; vaccine;
KW   breast cancer; B18Aq1; B11Aq1; B15Aq1; ss.
XX
OS   Homo sapiens.
XX
PN   WO200061753-A2.
XX
PD   19-OCT-2000.
XX
PF   07-APR-2000; 2000WO-US09312.
XX
PR   09-APR-1999;   99US-0289198.
PR   28-OCT-1999;   99US-0429755.
PR   23-MAR-2000;   2000US-0534825.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Prudakis TN, Smith JM, Reed SG, Misher LE, Retter MW, Dillon DC;
XX
DR   WPI; 2000-628403/60.
DR   P-PSDB; AAB28630.
XX
PT   An isolated polypeptide comprising an immunogenic portion of a breast
PT   tumor protein used for inhibiting the development of cancer, especially
PT   breast cancer, and monitoring cancer progression in a patient
XX
PS   Claim 4; Page 178; 187pp; English.

XX
CC   The present sequence is given in a specification relating to compositions
CC   and methods for the treatment and diagnosis of breast cancer. Nucleotide
CC   sequences that are preferentially expressed in breast tumour tissue, and
CC   the polypeptides encoded by such nucleotide sequences, are used in
CC   compositions and vaccines to inhibit the development of cancer,
CC   especially breast cancer. The progression of a cancer may be monitored by
CC   carrying out detection of tumour specific antigens at subsequent time
CC   points and comparing the results from the different time points.
CC   CD4+ and/or CD8- T-Cells isolated from the cancer patient may be treated
CC   with tumour-specific polypeptides, polynucleotides encoding the
CC   polypeptides or antigen presenting cells expressing the polypeptides. The
CC   cells are then administered to the patient to inhibit development of
CC   cancer.
XX
SQ   Sequence 2040 BP; 716 A; 392 C; 500 G; 432 T; 0 other;


  Query Match             98.0%;  Score 1131.6;  DB 21;  Length 2040;
  Best Local Similarity;  99.6%;  Pred. No. 0;
  Matches 1134;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 atggtggttgaggttgattccatgccggctgcctcttctgtgaaagaagccatttggtctc 60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 atggtggttgaggttgattccatgccggctgcctcttctgtgaaagaagccatttggtctc 60

Qy      61 aggagcaagatgggcaagtggtgctgccgttgcttccccctgctgcagggaagcaggcaag 120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 aggagcaagatgggcaagtggtgctgccgttgcttccccctgctgcagggaagcaggcaag 120

Qy     121 atgaatgtgggcactttctggaagaccacgacgactctgctatgaaagacactcaggaagcaag 180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 atgaatgtgggcactttctggaagaccacgacgactctgctatgaaagacactcaggaagcaag 180

Qy     181 atgggcaagtggtgccgccactgcttccccctgctgcagggggaagtggcaagaagcaagtg 240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 atgggcaagtggtgccgccactgcttccccctgctgcagggggaagtggcaagaagcaagtg 240

Qy     241 ggcagcttctgggagaccacgacgactctgctatgaaagacactcaggaagcaagatgggcaag 300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 ggcagcttctgggagaccacgacgactctgctatgaaagacactcaggaagcaagatgggcaag 300

Qy     301 tggtggttgccgctgcttccccctgctgcagggggaagcggcaagaagcaagatgggcagctta 360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 tggtggttgccgctgcttccccctgctgcagggggaagcggcaagaagcaagatgggcagctta 360

Qy     361 ggaggactacgatgacaagtgccttcatggaagcccaggtaccacgtccgtgggaggaagatct 420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 ggaggactacgatgacaagtgccttcatggaagcccaggtaccacgtccgtgggaggaagatct 420

Qy     421 gacaagctccacagaggctgccttggtgggtaaagtccccagaaaaggatctcatcgtcatg 480
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 gacaagctccacagaggctgccttggtgggtaaagtccccagaaaaggatctcatcgtcatg 480

Qy     481 ctccagggacactgacgtgaacaagaaggacaagcaaaagaggacttgctctacatctggcc 540
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 ctccagggacactgacgtgaacaagaaggacaagcaaaagaggacttgctctacatctggcc 540

Qy     541 tctgccaatgggaattcagaagtagtaaaactcctgctgaacagacgatgtcaacttaat 600
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 tctgccaatgggaattcagaagtagtaaaactcctgctgaacagacgatgtcaacttaat 600

Qy     601 gtccttgacaacaaaaagagggacagctctgataaaggccgtacaatgccagggaagatgaa 660
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 gtccttgacaacaaaaagagggacagctctgataaaggccgtacaatgccagggaagatgaa 660

Qy     661 tgtgcgttaatgttgctggaacatggcaattgatctaaaatattccagatgagtatggaaat 720
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 tgtgcgttaatgttgctggaacatggcaattgatctaaaatattccagatgagtatggaaat 720

Qy     721 accactctgcactacgctatctataatgaaggataaattaattgtgccaaagcaattgctctta 780
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 accactctgcactacgctatctataatgaaggataaattaattgtgccaaagcaattgctctta 780
```



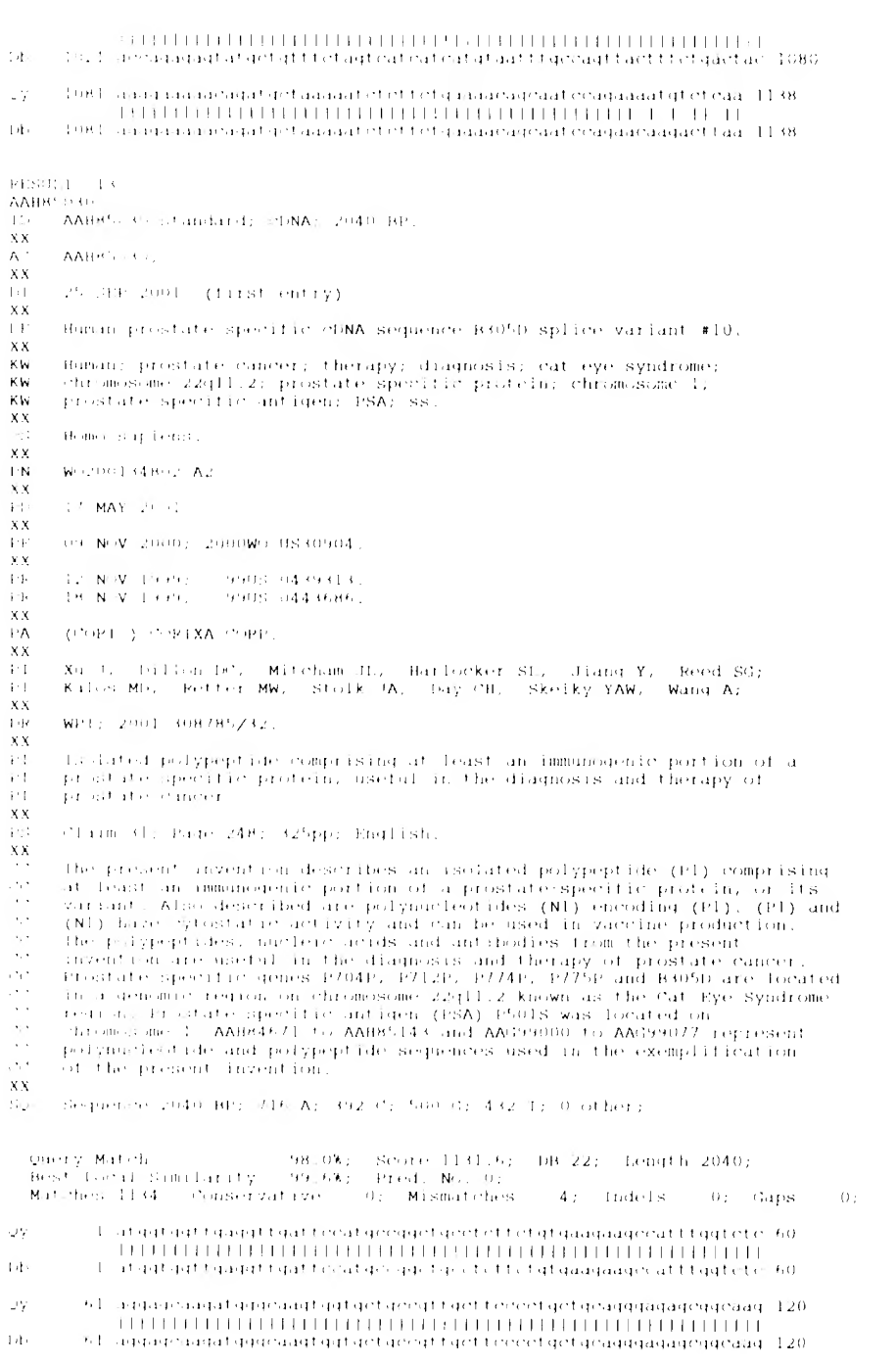
RESULT 11
AAA06608
ID AAA06608 standard; cDNA; 2040 BP.
XX
AC AAA06608;
XX
DT 24-JUN-2000 (first entry)
XX
DE Human immunogenic prostate tumour protein cDNA sequence SEQ ID NO:675.
XX
KW Human; prostate cancer; diagnosis; tumour; gene therapy; detection;
KW immunotherapy; cytostatic; vaccine; ss.
XX
OS Homo sapiens.
XX
PN WO200044895-A2.
XX
PD 27-JAN-2000.
XX
PF 24-JUL-1999; 99WO-US15838.
XX
PR 24-JUL-1998; 98US-0115454.
PR 24-JUL-1998; 98US-0115434.
PR 25-SEP-1998; 98US-0159812.
PR 25-SEP-1998; 98US-0159822.
PR 15-JAN-1999; 99US-0232149.
PR 15-JAN-1999; 99US-0232080.
PR 09-APR-1999; 99US-0288946.
XX
PA (CORI-) CORIXA CORP.
XX
PI Dillon DC; Harlocker SL; Yoqiu J; Xu J; Mitcham JL;
XX
DR WPI; 2000-171268/15.
XX
PT New polypeptide useful for treating and diagnosing prostate cancer
PT comprises an immunogenic portion of prostate tumor protein
XX
PS Claim 50; Page 225-224; 264pp; English.
XX
CC The present invention describes isolated polypeptides, comprising an
CC immunogenic portion of a prostate tumour protein (PTP). The polypeptides
CC and polynucleotides encoding them have cytostatic activity and can be
CC used in vaccines and in gene therapy. The polypeptides and
CC polynucleotides encoding them, antigen presenting cells which express
CC the polypeptides, antibodies against the polypeptides and vaccines
CC comprising them can be used for inhibiting the development of prostate
CC cancer in a patient. The polypeptides can be used to generate antibodies
CC or anti-idiotypic antibodies for passive immunotherapy. A portion of
CC the polynucleotides encoding the polypeptides can be used as a probe or
CC to modulate the expression of the polypeptides. AAA06611 to AAA06631 and
CC AAY82000 to AAY82020 represent sequences used in the exemplification of
CC the present invention.
XX
SQ Sequence 2040 BP; 716 A; 493 C; 500 G; 431 T; 0 other

Query Match 98.0%; Score 1151.6; DB 21; Length 2040;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 1134; Conservative 0; Mismatches 4; Indels 0; Gaps 0;



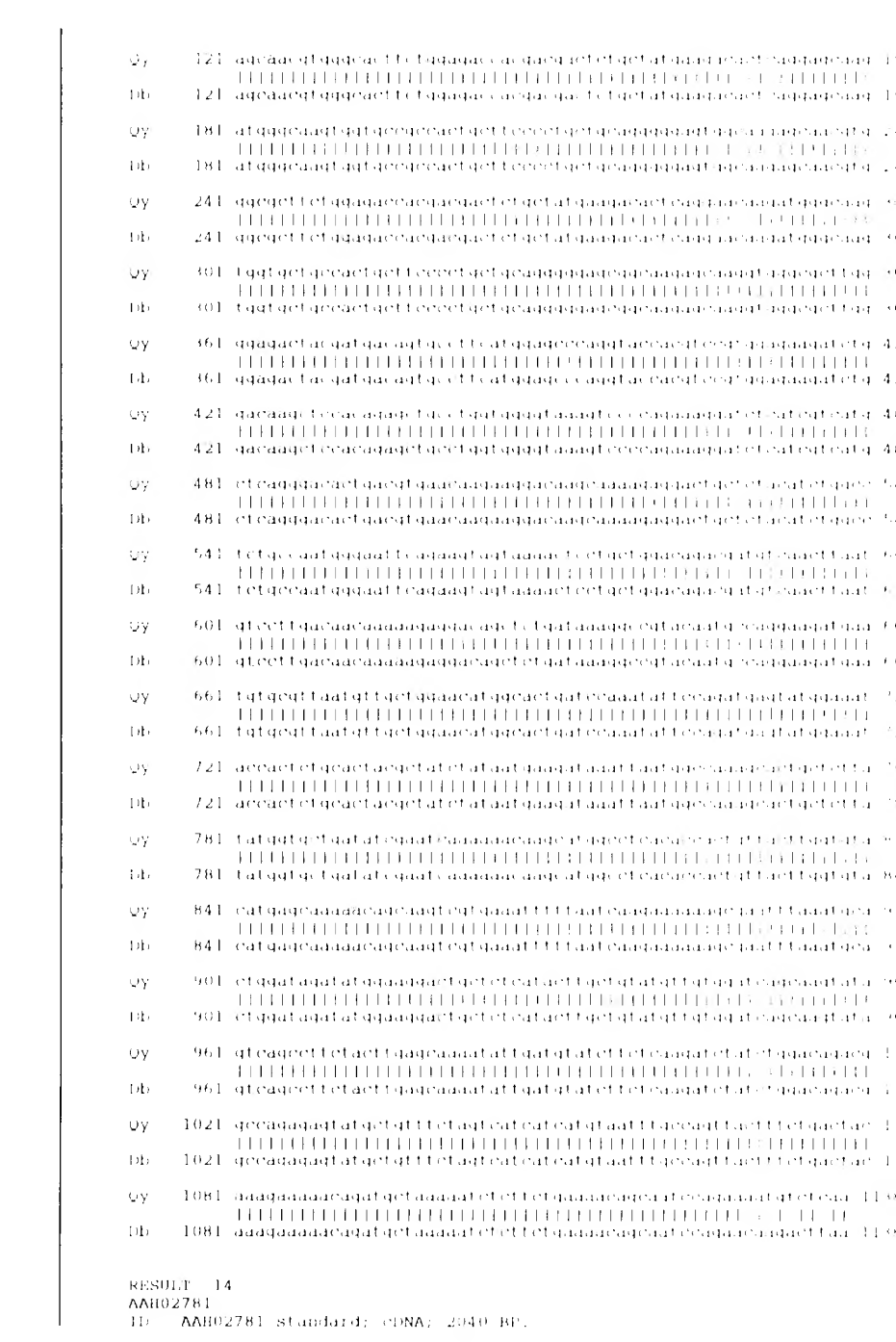
```
Qy     901 ctggataqatatqqaaqqactqctctcatacttqctqtatqttqtqqatcaqcaaqtata 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 ctggataqatatqqaaqqactqctctcatacttqctqtatqttqtqqatcaqcaaqtata 960

Qy     961 qtcaqccttctacttqaqcaaaatattqatqtatcttctcaaqatctatctqqacaqacq 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 qtcaqccttctacttqaqcaaaatattqatqtatcttctcaaqatctatctqqacaqacq 1020

Qy    1021 qccaqaqaqtatqctqtttctaqtcatcatcatqtaatttqcccaqttacttttctqactac 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 qccaqaqaqtatqctqtttctaqtcatcatcatqtaatttqcccaqttacttttctqactac 1080

Qy    1081 aaaqaaaaaacaqatqctaaaaatctcttctqaaaacaqcaatccaqaaaatqtctcaa 1138
           |||||||||||||||||||||||||||||||||||||||||||||||| | | || ||
Db    1081 aaaqaaaaaacaqatqctaaaaatctcttctqaaaacaqcaatccaqaaacaaqacttaa 1138


RESULT 12
AAH93716
ID   AAH93716 standard; cDNA; 2040 BP.
XX
AC   AAH93716;
XX
DT   04-OCT-2001 (first entry)
XX
DE   Human prostate specific cDNA sequence B305D splice variant #10.
XX
KW   Human; prostate cancer; prostate-specific; diagnosis; vaccine;
KW   cytostatic; gene therapy; metastasis; ss.
XX
OS   Homo sapiens.
XX
PN   WO200151633-A2.
XX
PD   19-JUL-2001.
XX
PF   16 JAN-2001; 2001WO-US01574.
XX
PR   14-JAN-2000; 2000US-0484672.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG;
PI   Kalos MD, Panzer CR, Day CH, Retter MW, Stolk JA, Skeiky YAW,
PI   Wang A, Meagher MJ;
XX
DR   WPI; 2001-425873/45.
XX
PT   New polynucleotide encoding a prostate-specific protein, for
PT   diagnosing, monitoring and treating prostate cancer in a patient and
PT   for use in vaccines -
XX
PS   Claim 1; Page 348; 543pp; English.
XX
CC   The present invention describes polynucleotide sequences (I) which encode
CC   prostate-specific proteins (II). (I) and (II) have cytostatic activity,
CC   and can be used in vaccine production and gene therapy. (I), (II),
CC   antibodies to (II), fusion proteins comprising (II), and isolated
CC   T cells prepared using (I) or (II) are used treat cancer in a patient.
CC   (I) and the antibodies are also used in the detection of cancer in a
CC   patient. The cancer that is diagnosed or treated is particularly
CC   prostate cancer. (I) and (II) can be used in vaccines. The antibodies or
CC   (I) can be used for monitoring the progression of cancer in a patient.
CC   (I) and (II) can also be used to improve diagnostic and therapeutic
CC   methods for prostate cancer. They can indicate the level of metastasis
CC   as well as the prostate volume. AAH93357 to AAH93944 and AAM01115 to
CC   AAM01318 represent polynucleotide and amino acid sequences used in the
CC   exemplification of the present invention.
XX
SQ   Sequence 2040 BP; 716 A; 392 C; 509 G; 432 T; 0 other;
```

```
  Query Match             98.0%;  Score 1131.6;  DB 22;  Length 2040;
  Best Local Similarity   99.6%;  Pred. No. 0;
  Matches 1134;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 atggtggttgaggttgattccatgccggctgcctcttctgtgaaqaaqccatttggtctc 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 atggtggttgaggttgattccatgccggctgcctcttctgtgaaqaaqccatttggtctc 60

Qy      61 aggaqcaaqatqqqcaaqtqqtqctqccqttqcttcccctqctqcaqqqaqaqcqqcaaq 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 aggaqcaaqatqqqcaaqtqqtqctqccqttqcttcccctqctqcaqqqaqaqcqqcaaq 120

Qy     121 aqcaacqtqqqcacttctqqaqaccacqacqactctqctatqaaqacactcaqcaqcaaq 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 aqcaacqtqqqcacttctqqaqaccacqacqactctqctatqaaqacactcaqcaqcaaq 180

Qy     181 atqqqcaaqtqqtqccqccactqcttcccctqctqcaqqqqqaqtqqcaaqaqcqaccta 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 atqqqcaaqtqqtqccqccactqcttcccctqctqcaqqqqqaqtqqcaaqaqcqaccta 240

Qy     241 qqcqcttctqaqqaqccacqacqactctqctatqaaqacactcaqqaqcaaqatcqqcaa 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 qqcqcttctqaqqaqccacqacqactctqctatqaaqacactcaqqaqcaaqatcqqcaa 300

Qy     301 tqqtqctqccactqcttcccctqctqcaqqqqqaqcqqcaaqaqcaaqqtqqqcqcttqa 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 tqqtqctqccactqcttcccctqctqcaqqqqqaqcqqcaaqaqcaaqqtqqqcqcttqa 360

Qy     361 qqaqactacqatqacaatqcttctatqqaqccaqqtaccacqtccqtqqaqaaqatctta 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 qqaqactacqatqacaatqcttctatqqaqccaqqtaccacqtccqtqqaqaaqatctta 420

Qy     421 qacaaqctccacaqaqctqcctqqtqqqqtaaaqtcccccaqaaaaqqatctcatcqtcata 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 qacaaqctccacaqaqctqcctqqtqqqqtaaaqtcccccaqaaaaqqatctcatcqtcata 480

Qy     481 ctcaqqqacactqacqtqaacaaqaaqqacaaqcaaaaqaqqactqctctacatctqccc 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 ctcaqqqacactqacqtqaacaaqaaqqacaaqcaaaaqaqqactqctctacatctqccc 540

Qy     541 tctqccaatqqqaattcaqaaqtaqtaaaactcctqctqqacaqacqatqtcaacttaat 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 tctqccaatqqqaattcaqaaqtaqtaaaactcctqctqqacaqacqatqtcaacttaat 600

Qy     601 qtcttqacaacaaaaaqaqqacaqctctqatqaaqqccqtacaatqccaqqaaqatqaa 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 qtcttqacaacaaaaaqaqqacaqctctqatqaaqqccqtacaatqccaqqaaqatqaa 660

Qy     661 tqtqcqttaatqttqctqqaacatqqcactqatccaaatattccaqatqaqtatqqaaat 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 tqtqcqttaatqttqctqqaacatqqcactqatccaaatattccaqatqaqtatqqaaat 720

Qy     721 accactctqcactacqctatctataatqaaqataaattaatqqccaaaqcactqctctta 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 accactctqcactacqctatctataatqaaqataaattaatqqccaaaqcactqctctta 780

Qy     781 tatqqtqctqatatcqaatcaaaaaacaaqcatqqcctcacaccactqttacttqtqta 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 tatqqtqctqatatcqaatcaaaaaacaaqcatqqcctcacaccactqttacttqtqta 840

Qy     841 catqaqcaaaaacaqcaaqtcqtqaaatttttaatcaaqaaaaaaqcqaatttaaatqca 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 catqaqcaaaaacaqcaaqtcqtqaaatttttaatcaaqaaaaaaqcqaatttaaatqca 900

Qy     901 ctqqataqatatqqaaqqactqctctcatacttqctqtatqttqtqqatcaqcaaqtata 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 ctqqataqatatqqaaqqactqctctcatacttqctqtatqttqtqqatcaqcaaqtata 960

Qy     961 qtcaqccttctacttqaqcaaaatattqatqtatcttctcaaqatctatctqqacaqacq 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 qtcaqccttctacttqaqcaaaatattqatqtatcttctcaaqatctatctqqacaqacq 1020

Qy    1021 qccaqaqaqtatqctqtttctaqtcatcatcatqtaatttqcccaqttacttttctqactac 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 qccaqaqaqtatqctqtttctaqtcatcatcatqtaatttqcccaqttacttttctqactac 1080
```



RESULT 13
AAH80940
ID AAH80940 standard; cDNA; 2040 BP.
XX
AC AAH80940;
XX
DT 25-SEP-2001 (first entry)
XX
DE Human prostate specific cDNA sequence B305D splice variant #10.
XX
KW Human; prostate cancer; therapy; diagnosis; cat eye syndrome;
KW chromosome 22q11.2; prostate specific protein; chromosome 1;
KW prostate specific antigen; PSA; ss.
XX
OS Homo sapiens.
XX
PN WO200134802-A2.
XX
PD 17-MAY-2001.
XX
PF 09-NOV-2000; 2000WO-US30904.
XX
PR 12-NOV-1999; 99US-0439413.
PR 18-NOV-1999; 99US-0443686.
XX
PA (PROP) PROMIXA PROP.
XX
PI Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG;
PI Kales MD, Potter MW, Stolk JA, Day CH, Skelky YAW, Wang A;
XX
DR WPI; 2001-308785/32.
XX
PT Isolated polypeptide comprising at least an immunogenic portion of a
PT prostate specific protein, useful in the diagnosis and therapy of
PT prostate cancer
XX
PS Claim 61; Page 248; 325pp; English.
XX
CC The present invention describes an isolated polypeptide (P1) comprising
CC at least an immunogenic portion of a prostate-specific protein, or its
CC variant. Also described are polynucleotides (N1) encoding (P1), (P1) and
CC (N1) have cytostatic activity and can be used in vaccine production.
CC The polypeptides, nucleic acids and antibodies from the present
CC invention are useful in the diagnosis and therapy of prostate cancer.
CC Prostate specific genes B704P, P712P, P774P, P775P and B305D are located
CC in a genomic region on chromosome 22q11.2 known as the Cat Eye Syndrome
CC region. Prostate specific antigen (PSA) P501S was located on
CC chromosome 1. AAH84471 to AAH85143 and AAG99000 to AAG99077 represent
CC polynucleotide and polypeptide sequences used in the exemplification
CC of the present invention.
XX
SQ Sequence 2040 BP; 716 A; 392 C; 500 G; 432 T; 0 other;

Query Match 98.0%; Score 1141.6; DB 22; Length 2040;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1134; Conservative 0; Mismatches 4; Indels 0; Gaps 0;



RESULT 14
AAH02781
ID AAH02781 standard; cDNA; 2040 BP.

```
XX
AC   AAH02781;
XX
DT   14-JUN-2001  (first entry)
XX
DE   Prostate tumour antigen determined cDNA splice variant of B305D #10.
XX
KW   Human; prostate tumour antigen; prostate tumour; therapy; diagnosis;
KW   prostate cancer; immunogenic; cytostatic; vaccine; ss.
XX
OS   Homo sapiens.
XX
PN   WO200125272 A2.
XX
PD   12-APR-2001.
XX
PF   04-OCT-2000; 2000WO-US27464
XX
PR   04-OCT-1999;   99US-0157455.
XX
PA   (CORI-) CORIXA CORP
XX
PI   Xu J,  Skeiky YAW,  Reed SG,  Cheever MA;
XX
DR   WPI; 2001-215962/22.
DR   P-PSDB; AAB74817.
XX
PT   Prostate specific protein and its encoding polynucleotide, useful for
PT   the treatment and diagnosis of prostate cancer
XX
PS   Claim 50; Page 233; 276pp; English.
XX
CC   The present invention describes an isolated polypeptide (I) comprising
CC   at least an immunogenic portion of a prostate tumour antigen protein or
CC   its variant. (I) have cytostatic activity and can be used in vaccine
CC   production. (I), prostate tumour antigen polynucleotides, an antigen
CC   presenting cell (APC e.g. a dendritic cell) that expresses (I), and a
CC   pharmaceutical composition containing (I) are useful for inhibiting the
CC   development of cancer in a patient. Antibodies specific for prostate
CC   specific proteins and oligonucleotides that hybridise to a
CC   polynucleotide that encodes a prostate specific protein are useful
CC   for detecting the presence or absence of a cancer or monitoring the
CC   progression the progression of a cancer, especially prostate cancer.
CC   AAH02422 to AAH02872, AAB74798 to AAB74821 and AAB74830 are sequences
CC   used in the exemplification of the present invention.
XX
SQ   Sequence 2040 BP; 716 A; 392 C; 508 G; 432 T; 0 other;
```

```
  Query Match           98.0%;  Score 1141.6;  DB 22;  Length 2040;
  Best Local Similarity  99.6%;  Pred. No. 0;
  Matches 1134;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 atggtggttgaggttgattccatgccggctgcctcttctgtgaagaagccatttggtctc 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 atggtggttgaggttgattccatgccggctgcctcttctgtgaagaagccatttggtctc 60

Qy      61 aggaggcaagatgggcaagtggtgctgccgttgcttcccctgctgcagggaggaggcggcaag 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 aggaggcaagatgggcaagtggtgctgccgttgcttcccctgctgcagggaggaggcggcaag 120

Qy     121 aggaacgtgggcacttctggaggaccacgacgacgactctgctatgaagacactcaggaggcaag 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 aggaacgtgggcacttctggaggaccacgacgacgactctgctatgaagacactcaggaggcaag 180

Qy     181 atgtgcaagtggtgctgccgccactgcttcccctgctgcagggaggaggcggcaaggtg 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 atgggcaagtggtgccgccactgcttcccctgctgcagggaggaggcggcaaggtg 240

Qy     241 ggccgcttctggaggaccacgacgacgactctgctatgaagacactcaggaacaagatgggcaag 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 ggccgcttctggaggaccacgacgacgactctgctatgaagacactcaggaacaagatgggcaag 300

Qy     301 tggtgctgccgccactgcttcccctgctgcagggaggaggcggcaaggaggcaaggtgggccgcttgg 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 tggtgctgccgccactgcttcccctgctgcagggaggaggcggcaaggaggcaaggtgggccgcttgg 360

Qy     361 ggaggactacgatgacgatgccttcatggaggccgaggtaccacgtccgtggaggaagatctg 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 ggaggactacgatgacgatgccttcatggaggccgaggtaccacgtccgtggaggaagatctg 420

Qy     421 gacgaggctccacagaggctgccttggtgggggtaaagtccccgaggaaggaggatctcatggtcatg 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 gacgaggctccacagaggctgccttggtgggggtaaagtccccgaggaaggaggatctcatggtcatg 480

Qy     481 ctccagggacactgacgtggacgaggaggacgaggcggaaggaggactgctctgtcatgtctggcc 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 ctccagggacactgacgtggacgaggaggacgaggcggaaggaggactgctctgtcatgtctggcc 540

Qy     541 tctgccaatgggaattcagaagtagtaaaactcctgctgggacgagcgatgtccaacttaat 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 tctgccaatgggaattcagaagtagtaaaactcctgctgggacgagcgatgtccaacttaat 600

Qy     601 gtcctggacgacgaggaggagggacgagctctgatgaaggccgtacaatgccgaggaagatgaa 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 gtcctggacgacgaggaggagggacgagctctgatgaaggccgtacaatgccgaggaagatgaa 660

Qy     661 tgtgcgttaatgttgctgcgaacatgcgcactgatccaaatattccagatgagtatggaaat 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 tgtgcgttaatgttgctgcgaacatgcgcactgatccaaatattccagatgagtatggaaat 72

Qy     721 acgactctgcactacgctatctataatgaagatgaattaatgccgaaagccctactcttta 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 acgactctgcactacgctatctataatgaagatgaattaatgccgaaagccctactcttta 780

Qy     781 tatggtgctgatatcgaatcaaaaaacaagcatggctgccacacactgttacttggtgtta 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 tatggtgctgatatcgaatcaaaaaacaagcatggctgccacacactgttacttggtgtta 840

Qy     841 catgaggcaaaaacagcaagtcgtgaaatttttaatcaagaaaaaagcgaatttaaatgca 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 catgaggcaaaaacagcaagtcgtgaaatttttaatcaagaaaaaagcgaatttaaatgca 900

Qy     901 ctggatagatatgaaaggactgctctcatacttgctgtatgttgtggatcaggcaagtata 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 ctggatagatatgaaaggactgctctcatacttgctgtatgttgtggatcaggcaagtata 960

Qy     961 gtcaggcttctacttgaggcaaaatattgatgtatcttctcaaaatctatctggacaggacg 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 gtcaggcttctacttgaggcaaaatattgatgtatcttctcaaaatctatctggacaggacg 1020

Qy    1021 gccaggaggtatgctgtttctagtcatcatcatgtaatttgccagttacttttctgactac 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 gccaggaggtatgctgtttctagtcatcatcatgtaatttgccagttacttttctgactac 1080

Qy    1081 aaaaaaaaaacagatgctaaaaatctcttctgaaaacagcaatccaggaaaatgtctcaa 1138
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 aaaaaaaaaacagatgctaaaaatctcttctgaaaacagcaatccaggaaaatgtctcaa 1138
```

```
RESULT  15
AAA06599
ID   AAA06599 standard; cDNA; 2000 BP.
XX
AC   AAA06599;
XX
DT   13 JUN 2000  (first entry)
XX
DE   Human immunogenic prostate tumour protein cDNA sequence SEQ ID NO:374.
XX
KW   Human; prostate cancer; diagnosis; tumour; gene therapy; detection;
KW   immunogenic; cytostatic; vaccine; ss.
XX
OS   Homo sapiens.
XX
PN   WO200004149-A2.
```

```
XX
PD   [illegible] JAN 2000.
XX
PF   14 [illegible] 1999;   99WO-US15838.
XX
PR   14 JUL 1998;   98US-0115453.
PR   14 JUL 1998;   98US-0116134.
PR   23 SEP 1998;   98US-0159812.
PR   23 SEP 1998;   98US-0159822.
PR   [illegible] JAN 1999;   99US-0232149.
PR   [illegible] JAN 1999;   99US-0232880.
PR   [illegible] APR 1999;   99US-0288946.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Dillon DC, Harlocker SL, Yuqiu J, Xu J, Mitcham JL;
XX
DR   WPI; 2000-[illegible]171268/15.
XX
PT   New polypeptide useful for treating and diagnosing prostate cancer
PT   comprises an immunogenic portion of prostate tumor protein -
XX
PS   Claim 50; Page 222-223; 263pp; English.
XX
CC   The present invention describes isolated polypeptides, comprising an
CC   immunogenic portion of a prostate tumour protein (PTP). The polypeptides
CC   and polynucleotides encoding them have cytostatic activity and can be
CC   used in vaccines and in gene therapy. The polypeptides and
CC   polynucleotides encoding them, antigen presenting cells which express
CC   the polypeptides, antibodies against the polypeptides and vaccines
CC   comprising them can be used for inhibiting the development of prostate
CC   cancer in a patient. The polypeptides can be used to generate antibodies
CC   or anti-idiotypic antibodies for passive immuno therapy. A portion of
CC   the polynucleotides encoding the polypeptides can be used as a probe or
CC   to modulate the expression of the polypeptides. AAA06241 to AAA06691 and
CC   AAY8[illegible] to AAY82020 represent sequences used in the exemplification of
CC   the present invention.
XX
SQ   Sequence 2000 BP; 598 A; 487 C; 489 G; 426 T; 0 other;

  Query Match             97.8%;  Score 1130;  DB 21;  Length 2000;
  Best Local Similarity   99.6%;  Pred. No. 0;
  Matches 1133;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 atgatggttgaggttgattccatgccggctgcctcttctgtgaagaagccatttggtctc 60
Db          1 atgatggttgaggttgattccatgccggctgcctcttctgtgaagaagccatttggtctc 60

Qy         61 aggagcaagatgggcaagtggtgctgccgttgcttcccctgctgcagggagagcggcaag 120
Db         61 aggagcaagatgggcaagtggtgctgccgttgcttcccctgctgcagggagagcggcaag 120

Qy        121 agcaacgtgggcacttctggagaccacgacgactctgctatgaagacactcaggagcaag 180
Db        121 agcaacgtgggcacttctggagaccacgacgactctgctatgaagacactcaggagcaag 180

Qy        181 atgggcaagtggtgcggcgcactgcttcccctgctgcagggggagtggcaagagcaacgtg 240
Db        181 atgggcaagtggtgcggcgcactgcttcccctgctgcagggggagtggcaagagcaacgtg 240

Qy        241 ggcgcttctggagaccacgacgactctgctatgaagacactcaggaacaagatgggcaag 300
Db        241 ggcgcttctggagaccacgacgactctgctatgaagacactcaggaacaagatgggcaag 300

Qy        301 tggtgctgccactgcttcccctgctgcagggggagcggcaagagcaagtggggcgcttgg 360
Db        301 tggtgctgccactgcttcccctgctgcagggggagcggcaagagcaagtggggcgcttgg 360

Qy        361 ggaactacgatgacagtgccttcatggagcccaggtaccacgtccgtggagaagatctg 420
Db        361 ggaactacgatgacagtgccttcatggagcccaggtaccacgtccgtggagaagatctg 420

Qy        421 gacaagctccacagagctgcctgatggggtaaagtccccagaaaggatttcatcgtcatg 480
Db        421 gacaagctccacagagctgcctgatggggtaaagtccccagaaaggatttcatcgtcatg 480

Qy        481 ctcagggacactgacgtgaacaagaaggacaagcaaaagaagactgctctacatctggcc 540
Db        481 ctcagggacactgacgtgaacaagaaggacaagcaaaagaagactgctctacatctggcc 540

Qy        541 tctgccaatgggaattcagaagtagtaaaactcctgctggacagacgatgtcaacttaat [illegible]
Db        541 tctgccaatgggaattcagaagtagtaaaactcctgctggacagacgatgtcaacttaat [illegible]

Qy        601 gtccttgacaacaaaaagaggacagctctgataaaggccgtacaatgccaggaagatgaa [illegible]
Db        601 gtccttgacaacaaaaagaggacagctctgataaaggccgtacaatgccaggaagatgaa [illegible]

Qy        661 tgtgcgttaatgttgctggaacatggcactgatccaaatattccagatgagtatggaaat [illegible]
Db        661 tgtgcgttaatgttgctggaacatggcactgatccaaatattccagatgagtatggaaat [illegible]

Qy        721 accactctgcactacgctatctataatgaagataaattaatggccaaagcactgctctta 780
Db        721 accactctgcactacgctatctataatgaagataaattaatggccaaagcactgctctta 780

Qy        781 tatggtgctgatatcgaatcaaaaaacaagcatggcctcacaccactgttactttggtgta 840
Db        781 tatggtgctgatatcgaatcaaaaaacaagcatggcctcacaccactgttactttggtgta 840

Qy        841 catgagcaaaaacagcaagtcgtgaaatttttaatcaagaaaaaagcgaatttaaatgca 900
Db        841 catgagcaaaaacagcaagtcgtgaaatttttaatcaagaaaaaagcgaatttaaatgca 900

Qy        901 ctggatagatatggaaggactgctctcatacttgctgtatgttgtggatcagcaagtata 960
Db        901 ctggatagatatggaaggactgctctcatacttgctgtatgttgtggatcagcaagtata 960

Qy        961 gtcagccttctacttgagcaaaatattgatgtatcttctcaagatctatctggacagacg 1020
Db        961 gtcagccttctacttgagcaaaatattgatgtatcttctcaagatctatctggacagacg 1020

Qy       1021 gccagagagtatgctgtttctagtcatcatcatgtaatttgccagttactttctgactac 1080
Db       1021 gccagagagtatgctgtttctagtcatcatcatgtaatttgccagttactttctgactac 1080

Qy       1081 aaagaaaaacagatgctaaaaatctcttctgaaaacagcaatccagaaaatgtctcaa 11[illegible]
Db       1081 aaagaaaaacagatgctaaaaatctcttctgaaaacagcaatccagaacagacttaa 11[illegible]
```

Search completed: March 9, 2002, 03:09:29
Job time: 6197 sec

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: March 9, 2002, 00:28:41 ; Search time 1484.81 Seconds
(without alignments)
8858.897 Million cell updates/sec

Title: US-09-699-295-301
Perfect score: 1155
Sequence: 1 atggtggttgaggttgatatc.....cdaaaaccaaaaataaaataa 1155

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 11451947 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estfun:*
2: em_esthum:*
3: em_estin:*
4: em_estom:*
5: em_estpl:*
6: em_estbac:*
7: em_estro:*
8: em_estvir:*
9: em_htc:*
10: gb_est1:*
11: gb_est2:*
12: gb_htc:*
13: gb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_rod:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No	Score	% Query Match	Length	DB	ID	Description
1	452.2	42.6	865	11	BF676987	BF676987 602084215
2	348.8	30.2	843	10	AU120666	AU120666 AU120666
3	272	23.5	549	10	AA910780	AA910780 o148b04.s
4	266.2	23.0	694	11	BG720647	BG720647 6026942528
5	251.2	21.7	1258	12	AK017783	AK017783 Mus muscu
6	217.6	18.8	747	11	BG717440	BG717440 602689791
7	217.6	18.8	1054	12	AK015948	AK015948 Mus muscu
8	213.6	18.5	822	11	BF675057	BF675057 602136651
9	211.6	18.3	810	11	BG703856	BG703856 602686889
10	209.6	18.1	907	11	BG721483	BG721483 602695061
11	207.8	18.0	916	11	BG716974	BG716974 602689005
12	207.4	18.0	1842	12	AK016475	AK016475 Mus muscu
13	206.2	17.9	1203	12	AK016868	AK016868 Mus muscu
14	206.2	17.9	1837	12	AK015247	AK015247 Mus muscu
15	205.2	17.8	1209	12	AK005925	AK005925 Mus muscu
c 16	202	17.5	380	11	BF329652	BF329652 RC5 BNGH
c 17	200.4	17.4	565	10	AI972614	AI972614 wt41a10.x
c 18	199.4	17.3	532	11	BF002867	BF002867 7q5c05.x
19	199	17.2	791	11	BG719053	BG719053 6026949174
20	197.6	17.1	1441	12	AK015473	AK015473 Mus muscu
21	197.6	17.1	1456	12	AK015012	AK015012 Mus muscu
22	197.4	17.1	809	11	BG717899	BG717899 602695548
c 23	196	17.0	461	10	AI424263	AI424263 tf0.c05.x
24	196	17.0	1400	12	AK015165	AK015165 Mus muscu
25	194	16.8	754	11	BG718834	BG718834 602697606
26	191.8	16.6	789	11	BG702026	BG702026 602668412
c 27	190.6	16.5	543	11	BF891930	BF891930 IL0 ML14
28	190.4	16.5	894	11	BF675049	BF675049 602136643
29	188.2	16.3	764	11	BG722165	BG722165 602721427
30	181.6	15.7	916	11	BG722538	BG722538 602720941
31	181.2	15.7	809	11	BF377398	BF377398 MR4 INCI1
32	177.2	15.3	792	11	BG200469	BG200469 BF11C302
33	176.6	15.3	717	11	BG720794	BG720794 602692216
34	174.8	15.1	531	13	AQ615477	AQ615477 BF 5144 B
35	174.8	15.1	786	11	BG217804	BG217804 RST42518
36	174	15.1	693	11	BG722705	BG722705 602694332
37	173	15.0	708	10	AW955793	AW955793 EST30 3663
38	172.8	15.0	646	11	BI105639	BI105639 602951701
39	168.2	14.6	461	10	AA164865	AA164865 zq41a01.s
c 40	167.4	14.5	503	10	AI675967	AI675967 we43d01.x
41	166.4	14.4	678	11	BG722960	BG722960 602722425
42	166	14.4	796	11	BG722676	BG722676 602720650
43	165.4	14.3	594	10	AV601969	AV601969 AV6 L969
44	165	14.3	553	10	BE750985	BE750985 20Z577 MA
45	162	14.0	455	11	BI025848	BI025848 CM3 M10 29

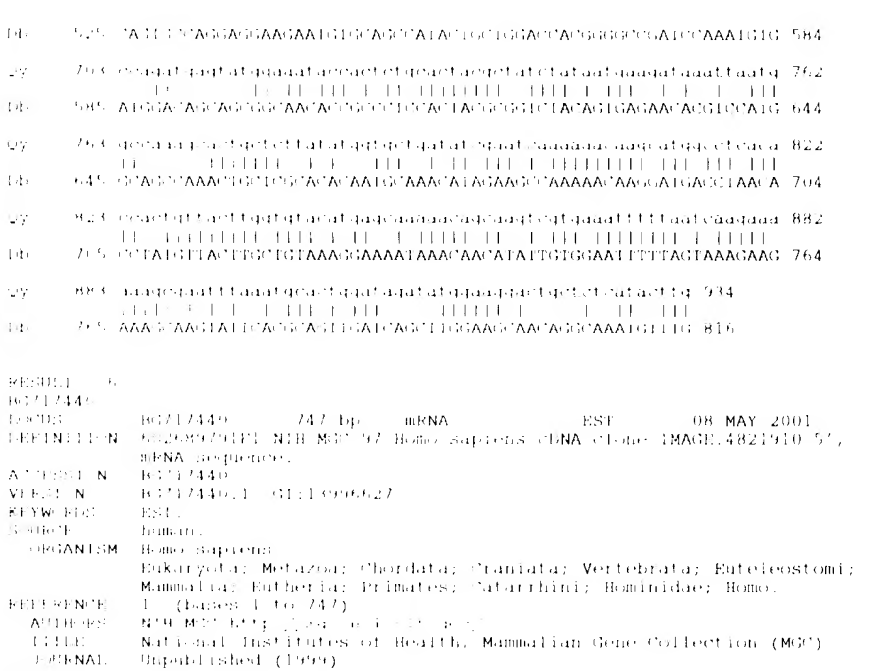
ALIGNMENTS

RESULT 1
BF676987
LOCUS BF676987 865 bp mRNA EST 21-DEC-2000
DEFINITION 602084215F1 NIH_MGC_83 Homo sapiens cDNA clone IMAGE:4248746 5', mRNA sequence.
ACCESSION BF676987
VERSION BF676987.1 GI:11950882
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 865)
AUTHORS NIH-MGC http://mgc.nci.nih.gov.
TITLE National Institutes of Health. Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: CLONETECH Laboratories, Inc.
cDNA Library Preparation: CLONETECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Plate: LLCM1067 row: m column: 03
High quality sequence stop: 642
FEATURES Location/Qualifiers
source 1..865
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4248746"
/clone_lib="NIH_MGC_83"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: prostate; Vector: pDNR-LIB (Clontech);

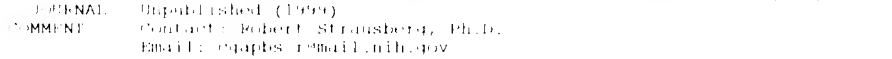
```
              Site_1: SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc
              ); 5' and 3' adaptors were used in cloning as follows: 5'
              adaptor sequence: 5'-CACGGCCATTATGGCC-3' and 3' adaptor
              sequence. 5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3'
              (where B = A, C, or G and N = A, C, G, or T). Average
              insert size 1.4 kb (range 0.5-4.0 kb). 14/15 colonies
              contained inserts by PCR. This library was enriched for
              full-length clones and was constructed by Clontech
              Laboratories (Palo Alto, CA)."
BASE COUNT      297 a    140 c    211 g    217 t
ORIGIN

  Query Match            42.6%;  Score 452.2;  DB 11;  Length 865;
  Best Local Similarity  98.3%;  Pred. No. 1.2e-108;
  Matches  508;  Conservative    0;  Mismatches    8;  Indels    1;  Gaps    1;

Qy       635 aggccgtacaatgccaggaagatgaatgtgcgttaatgttgctggaacatggcactgatc 694
Db        69 AGGCCGTACAATGCCAGGAAGATGAATGTCCGTTAATGTTGCTGGAACATGGCACTGATC 128

Qy       695 caaatattccagatgagtatggaaataccactctgcactacgctatctataatgaagata 754
Db       129 CAAACATTCCAGATGAGTATGGAAATACCACTCTACACTACGCTATCTACAATGAAGATA 188

Qy       755 aattaatggccaaagcactgctcttatatggtgctgatatcgaatcaaaaaacaagcatg 814
Db       189 AATTAATGGCCAAACCACTGCTCTTATACGGTGCTGATATCGAATCAAAAAACAAGCATG 248

Qy       815 gcctcacaccactgttacttggtgtacatgagcaaaaacagcaagtcgtgaaatttttaa 874
Db       249 GCCTCACACCACTGCTACTTGGTGTACATGACCAAAAACAGCAAGTGGTGAAATTTTTAA 308

Qy       875 tcaagaaaaaagcgaatttaaatgcactggatagatatggaaggactgctctcatacttg 934
Db       309 TCAAGAAAAAAGCGAATTTAAATGCCCTGGATAGATATGGAAGAACTGCTCTCATACTTG 368

Qy       935 ctgtatgttgtggatcagcaagtatagtcagccttctacttgagcaaaatattgatgtat 994
Db       369 CTGTATGTTGTGGATCAGCAAGTATAGTCAGCCTTCTACTTGAGCAAAATATTGATGTAT 428

Qy       995 cttctcaagatctatctggacagacggccagagagtatgctgtttctagtcatcatcatg 1054
Db       429 CTTCTCAAGATCTATCTGGACAGACGGCCAGAGAGTATGCTGTTTCTAGTCATCATCATG 488

Qy      1055 taatttgccagttactttctgactacaaagaaaaacagatgctaaaaatctcttctgaaa 1114
Db       489 TAA-TTGCCAGTTACTTTCTGACTACAAAGAAAAACAGATGCTAAAAATCTCTTCTGAAA 547

Qy      1115 acagcaatccagaaaatgtctcaagaaccagaaataa 1151
Db       548 ACAGCAATCCAGAAAATGTCTCAAGAACCAGAAATAA 584

RESULT 2
AU120666
LOCUS       AU120666       843 bp    mRNA            EST       19-OCT-2000
DEFINITION  AU120666 HEMBB1 Homo sapiens cDNA clone HEMBB1001175 5', mRNA
            sequence.
ACCESSION   AU120666
VERSION     AU120666.1  GI:10935901
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 843)
  AUTHORS   Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
            Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
            Isogai,T.
  TITLE     HRI human cDNA project
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Takao Isogai
            Genomics Laboratory
            Helix Research Institute
            1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            Tel: 81-438-52-3951
            Fax: 81-438-52-3952
            Email: genomics@hri.co.jp
            HRI human cDNA project; 5'- & 3'-end one pass sequencing. Helix
            Research Institute; cDNA library construction. Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            Helix Research Institute.
FEATURES             Location/Qualifiers
     source          1..843
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="HEMBB1001175"
                     /clone_lib="HEMBB1"
                     /tissue_type="whole embryo, mainly body"
                     /dev_stage="embryo, 10 weeks"
                     /note="Vector: pME18SFL3"
BASE COUNT      240 a    178 c    245 g    176 t      4 others
ORIGIN

  Query Match            30.2%;  Score 348.8;  DB 10;  Length 843;
  Best Local Similarity  69.2%;  Pred. No. 5.5e-74;
  Matches  492;  Conservative    0;  Mismatches  210;  Indels    9;  Gaps    1;

Qy       223 agtggcaagagcaacgtgggcgcttctggagaccacgacgactctgctatgaagacactc 282
Db       119 AGAGCCGAGAGAGACATCAGGTGTCTCTGAAGCCCGGTCGCCTGGGCCATGAAGAAGATT 178

Qy       283 aggaacaagatgggcaagtggtgctgccactgcttccccctgctgcaggggagcggcaag 342
Db       179 TTTAGTAAGAAGGGCGAGTCGCCCTTGGGCTCCTTCGCGCGGCGCCGGAGGAGCAGC    235

Qy       343 agcaaggtgggcgcttggggagactacgatgacagtgccttcatggagcccaggtaccac 402
Db       236 ---GCGGCAGCCGGGGCGAGCCGGGGAGGGCGCCTACTCGCAGCCCGGCTACCAC 289

Qy       403 gtccgtggagaagatctggacaagctccacagagctgcctggtggggtaaagtccccaga 462
Db       290 GTCCCAGACCGAGATCTCGGCAAGATCCACAAAGCTGCCAGCCGCGGTAATGTCCCGAAA 349

Qy       463 aaggatctcatcgtcatgctcagggacactgacgtgaacaagaaggacaagcaaaagaag 522
Db       350 GTGCAGCAGATCCTTTTGCTCAGGAAGAATCGCTTGAACGATAGACACAAGATGAACAGG 409

Qy       523 actgctctacatctggcctctgccaatgggaattcagaagtagtaaaactcctgctggac 582
Db       410 ACGGCTCTACATTTGGCCTGTGCCAATGGTCATCCAGAAGTAGTAACTCTCCTGGTGGAC 469

Qy       583 agacgatgtcaacttaatgtccttgacaacaaaaagaggacagctctgataaaggccgta 642
Db       470 AGAAAATGCCAGCTCAATGTCTGTGACAACGAAAACAGGACAGCTCTGATGAAGGCTCTA 529

Qy       643 caatgccaggaagatgaatgtgcgttaatgttgctggaacatggcactgatccaaatatt 702
Db       530 CAATGCCAGGAAGAGAAATGTGCAACTATTCTGCTAGAACATGGTGCTGATCCAAATCTT 589

Qy       703 ccagatgagtatggaaataccactctgcactacgctatctataatgaagataaattaatg 762
Db       590 GCGGATGTCCATGGCAACACTCCTCTTCACTATGCTGTCTATAATCAGGACATATCAGTA 649

Qy       763 gccaaagcactgctcttatatggtgctgatatcgaatcaaaaaacaagcatggcctcaca 822
Db       650 GCAACAAAGCTGCTTTTGTATGATGCAAATATTGAAGCAAAAAACAAGGATGACCTNACA 709

Qy       823 ccactgttacttggtgtacatgagcaaaaacagcaagtcgtgaaatttttaatcaagaaa 882
Db       710 CCACTTTTACTTGCAGTAAGTGGAAAAAAGCAGCAAATGGTGGAATTTTTAATAAAGAAA 769

Qy       883 aaagcgaatttaaatgcactggatagatatggaaggactgctctcatactt 933
```

```
Db        77 AAAGCAAATGTAAATGCCNTANATAAGTTGGAAAGCAGTCACCACTAATTT 820

RESULT      3
AA910780
LOCUS       AA910780        539 bp    mRNA            EST       10-JUN-1998
DEFINITION  ol48b04.s1 Soares NFL T GBC S1 Homo sapiens cDNA clone
            IMAGE:1526671 3' similar to TR:Q92527 Q92527 ANKYRIN MOTIF. ;, mRNA
            sequence.
ACCESSION   AA910780
VERSION     AA910780.1  GI:3050076
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 539)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 1548   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 235.
FEATURES             Location/Qualifiers
     source          1..539
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1526671"
                     /clone_lib="Soares NFL_T_GBC_S1"
                     /lab_host="DH10B"
                     /note="Organ: pooled; Vector: pT7T3D Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     Equal amounts of plasmid DNA from three normalized
                     libraries (fetal lung NbHL19W, testis NHT, and B-cell
                     NCI_CGAP_GCB1) were mixed, and ss circles were made in
                     vitro. Following HAP purification, this DNA was used as
                     tracer in a subtractive hybridization reaction. The driver
                     was PCR-amplified cDNAs from pools of 5,000 clones made
                     from the same 3 libraries. The pools consisted of
                     I.M.A.G.E. clones 297480-302087, 682632-687239,
                     726408-728711, and 729096-731399. Subtraction by Bento
                     Soares and M. Fatima Bonaldo. "
BASE COUNT      156 a    137 c    138 g    108 t
ORIGIN

  Query Match             23.5%;  Score 272;  DB 10;  Length 539;
  Best Local Similarity   73.4%;  Pred. No. 1.9e-55;
  Matches  361;  Conservative    0;  Mismatches  130;  Indels    1;  Gaps    1;

Qy        359 [illegible] 418
Db         41 GGGGCGAGCGGGGGAGGGCGCCTACTCGCAGCCGGCTACCACGTCCGAGACCGAGATC 100

Qy        419 [illegible] 478
Db        101 TGGGAAGATCCACAAAGCTGCCAGCGCGGTAATGTGGCGAAAGTGCAGCAGATCCTTT 160

Qy        479 [illegible]actgctctacatctgg 538
Db        161 TGCTCAGGAAGAATGGCTTGAACGATAGAGACAAGATGAACAGGACGGCTCTACATTTGG 220

Qy        539 [illegible]gctggacagacgatgtcaactta 598
Db        221 CCTGTGCCAATGGTCATCCAGAAGTAGTAACTCTCCTGGTGGACAGAAAATGCCAGCTCA 279

Qy        599 [illegible] 658
Db        280 ATGTCTGTGACAACGAAAACAGGACAGCTCTGATGAAGGCTGTACAAT[illegible]AGGAAGAGA [illegible]

Qy        659 aatgtgcgttaatgttgctggaacatggcactgatccaaatattccagatgagtatggaa 718
Db        340 AATGTGCAACTATTCTGCTAGAACATGGTGCTGATCAAATCT[illegible]ATGGCA [illegible]

Qy        719 ataccactctgcactacgctatctataatgaagataaattaatgg[illegible]actgctct 778
Db        400 ACACTGCTCTTCACTATGCTGTCTATAATGAGGACATATCAGTAGCAA[illegible]AAAGCTGCTTT [illegible]

Qy        779 tatatggtgctgatatcgaatcaaaaaacaagcatggcctcacaccactgttacttggtg 838
Db        460 TGTATGATGCAAATATTGAAGCACAAAACAGGATGACCTCACACCACT[illegible]TACTTGGTAG [illegible]

Qy        839 tacatgagcaaa 850
Db        520 TAAGTGGAAAAA 531

RESULT      4
BG720647
LOCUS       BG720647        694 bp    mRNA            EST       08-MAY-2001
DEFINITION  602692528F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4824630 5',
            mRNA sequence.
ACCESSION   BG720647
VERSION     BG720647.1  GI:13999834
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 694)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
             cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10736  row: h  column: 07
            High quality sequence stop: 694.
FEATURES             Location/Qualifiers
     source          1..694
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4824630"
                     /clone_lib="NIH_MGC_97"
                     /lab_host="DH10B"
                     /note="Organ: testis; Vector: pBluescriptR (modified
                     pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (dT-tag
                     ); oligo-dT primed using primer 5' [illegible]VN 3',
                     size-selected for average insert size 2.2 kb and
                     normalized to Rot 5. This is a primary library enriched
                     for full length clones and constructed using the
                     Cap-Trapper method (Carninci, in preparation). Library
                     constructed by M. Brownstein (NIMH/NHGRI, National
                     Institutes of Health). Note: this is a NIH MGC Library."
BASE COUNT      108 a    173 c    228 g    185 t
ORIGIN

  Query Match             23.0%;  Score 266.2;  DB 11;  Length 694;
  Best Local Similarity   89.7%;  Pred. No. 5e-54;
  Matches  286;  Conservative    0;  Mismatches   33;  Indels    0;  Gaps    0;

Qy          1 atggtagttgagattgattccatcggctgcctcttctgtgaaaaacccatttgatctc[illegible] [illegible]
```

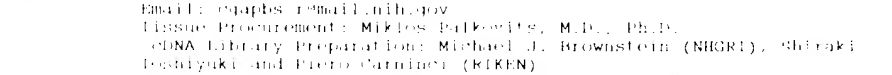
```
Db    376 ATGGTGGCTGAGGTTTGTTCAATGCCCGCTGCCTCTGCTGTCAACAAGCCATTTGATCTC 435

Qy     61 aggagcaagatgggcaagtggtgctgccgttgcttcccctgctgcaggggaggggcagg 120

Db    436 AGGAGCAAGATGGGCAAGTGGTGCCACCACCGCTTCCCCTGCTGCAGGGGAGGGGCAAG 495

Qy    121 agcaacgtgggcacttctggagaccacgacgactctgctatgaagacactcaggagcaag 180

Db    496 AGCAACATGGGCACTTCTGGAGACCACGACGACTCCTTTATGAAGACGCTCAGGAGCAAG 555

Qy    181 atgggcaagtggtgccgccactgcttcccctgctgcaggggagtggcaaggcagcgtg 240

Db    556 ATGGGCAAGTGTTGCCACCACTGCTTCCCTGCTGCAGGGGACGGCACCAGCAATCTG 615

Qy    241 ggcttctggagaccacgacgactctgctatgaagacactcaggaacaagatggcaag 300

Db    616 GGCACTTCTGCAGACCATGACAACTCCTTTATCAAGACACTCAGGAGCAAGATGGCAAG 675

Qy    301 tggtgctgccactgcttcc 319

Db    676 TGGTGCTGTCACTGTTCCC 694


RESULT      5
AK017783
LOCUS       AK017783     1758 bp    mRNA            HTC       05-JUL-2001
DEFINITION  Mus musculus 8 days embryo cDNA, RIKEN full-length enriched
            library, clone:5730521P14, full insert sequence.
ACCESSION   AK017783
VERSION     AK017783.1  GI:12857204
KEYWORDS    CAP trapper.
SOURCE      Mus musculus (strain:C57BL/6J) 8 days embryo cDNA to mRNA.
            clone_lib:RIKEN full-length enriched mouse cDNA library
            clone:5730521P14.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus
REFERENCE   1  (bases 1 to 1758)
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE     High-efficiency full length cDNA cloning
  JOURNAL   Methods in enzymology. 303, 19-44 (1999)
  MEDLINE   99279253
   PUBMED   10349636
REFERENCE   2  (bases 1 to 1758)
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL   Genome research. 10 (10), 1617-1630 (2000)
  MEDLINE   20499374
   PUBMED   11042159
REFERENCE   3  (bases 1 to 1758)
  AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
            Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
            Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
            Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
            Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
            Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
            Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
  TITLE     RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer
  JOURNAL   Genome research. 10 (11), 1757-1771 (2000)
  MEDLINE   20530913
   PUBMED   11076861
REFERENCE   4  (bases 1 to 1758)
  AUTHORS   The RIKEN Genome Exploration Research Group Phase II Team and the
            FANTOM Consortium.
  TITLE     Functional annotation of a full-length mouse cDNA collection
  JOURNAL   Nature 409, 685-690 (2001)
REFERENCE   5  (bases 1 to 1758)
  AUTHORS   Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A.,
            Arakawa,T., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M.,
            Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F.,
            Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kato,H., Kawai,J.,
            Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T.,
            Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y.,
            Okido,T., Owa,C., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H.,
            Sasaki,D., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T.,
            Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
            Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A.,
            Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
            Physical and Chemical Research (RIKEN), Laboratory for Genome
            Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
            RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
            Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
            Fax:81-45-503-9216)
COMMENT     Please visit our web site (http://genome.gsc.riken.go.jp/) for
            further details.
FEATURES             Location/Qualifiers
     source          1..1758
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /db_xref="MGD:MGI:1903054"
                     /db_xref="MGD:MGI:1917887"
                     /clone="5730521P14"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="8 days embryo"
     CDS             159..>1758
                     /note="putative"
                     /codon_start=1
                     /protein_id="BAB30930.1"
                     /db_xref="GI:12857205"
                     /translation="MKKIFGFFSKGPSPLGPSAFFFSNCVGFGRESASSHVPPYHIH
                     DKDMGKIHKAASVGDVAKVQHILILGKSGVNDRDKKDRTALHLACAYGHPEVVILLVE
                     KKCEIDARDSESSIALIKAVQCQEEECAAILLDHGADFNVMDSSGNIALHYAVYSENT
                     SMAAKLIAHNANIFAKNKI..IPMLLAVKENKQHIVEFLVKKKASIHAVDQLGSNRQM
                     FEYDGKRLQRSENSNPVDNGSEDGSLTRSYNTPGPADSWPTSDEEDYNFDNKNVPKIN
                     LTELWTAAQQSRKNQIKCGFEELDNGARFDDSDSPSESEDAIEVEPAPSVRVQILSPS
                     RQSPDPVEGATELAIEGEENGTDVIESASQEQPNHDNIPRADGWHKSNKSEMMSALGL
                     GEDEDEHSPWDSESISESVSLKDVGHFSGIADQIGKRRAHGQIEDVIYIPSCMSGSRS
                     FKMAKLEESRNVGLPVAHMGMEKNEAPRKYAIMEPTIERRAPVLNKIEIVGMIQAQTF
                     KSEPESVSREEQTRLSGSEDSQQRV"
BASE COUNT      531 a    392 c    474 g    361 t
ORIGIN

  Query Match             21.7%;  Score 251.2;  DB 12;  Length 1758;
  Best Local Similarity   64.0%;  Pred. No. 2.5e-50;
  Matches  379;  Conservative    0;  Mismatches  213;  Indels    0;  Gaps    0;

Qy    343 agcaaggtgggcgcttggggagactacgatgacagtgccttcatggagcccaggtaccac 402

Db    225 ACTAACTCCTCGGCTTCCACCTCACAGCCTACTCCCTCCCACCTGCCAGGTACCAC 284

Qy    403 gtccgtggagaagatctggacaagctccacagagctgcctggtaggggtaaagtcccaga 462

Db    285 ATTCACGACAAGGATATGGGCAAGATCCACAAGGCTGCAAGCGTGGGAGATGTAGCCAAA 344

Qy    463 aaggatctcatcgtcatgctcagggacactgacgtgaacaagaagacaagcaaaagacc 522

Db    345 GTGCAGCATATCCTTATCCTTGGAAAAAGTGGCGTGAACGATAGAGACAAGAAGGACAGG 404

Qy    523 actgctctacatctggcctctgccaatgggaattcagaagtagtaaaactcctgctggac 582

Db    405 ACTGCTCTACATCTTGCCTGTGCTTACGGCCACCCAGAAGTGGTGACTCTCTTAGTAGAG 464

Qy    583 agacgatgtcaacttaatgtccttgacaacaaaaagaggacagctctgataaagcccgta 642

Db    465 AGAAAATGTGAGATAGATGCCCGTGACAGCGAGAGCAGCACGGCCCTCATTAAGGCCGTG 524

Qy    643 caatgccaggaagatgaatgtgcgttaatgttgctggaacatggcactgatccaaatatt 702
```

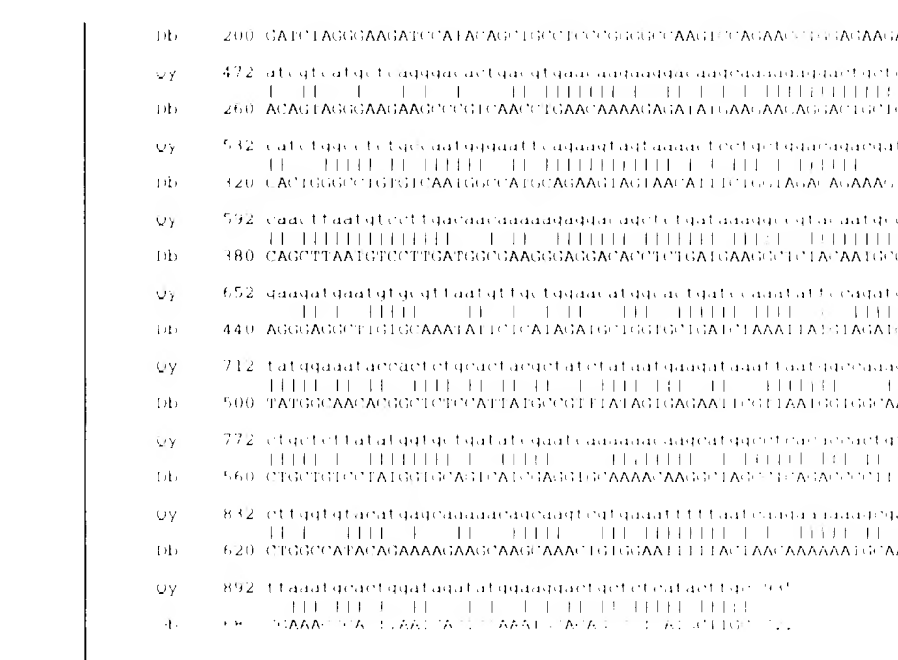


RESULT 6
BG717449
LOCUS BG717449 747 bp mRNA EST 08-MAY-2001
DEFINITION BG717449 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4821910 5', mRNA sequence.
ACCESSION BG717449
VERSION BG717449.1 GI:13096627
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 747)
AUTHORS NIH_MGC_97;
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://www.mcb.harvard.edu
Plate: LLAM10729 row: f column: 23
High quality sequence stop: 744.
FEATURES Location/Qualifiers
source 1..747
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4821910"
/clone_lib="NIH_MGC_97"
/lab_host="DH10B"
/note="Organ: testis; Vector: pBluescriptR (modified pBluescript KS+); Site 1: BamHI; Site 2: SalI-XbaI (stepwise); oligo dT primed using primer 5'-TTTTTTTTTTTTTTTVN-3', size selected for average insert size 2.2 kb and normalized to ROI 5. This is a primary library enriched for full length clones and constructed using the Cap trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIMH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT 229 a 163 c 225 g 130 t
ORIGIN



Query Match 18.8%; Score 217.6; DB 11; Length 747;
Best Local Similarity 66.0%; Pred. No. 2.9e-42;
Matches 344; Conservative 0; Mismatches 174; Indels 4; Gaps 2;





RESULT 7
AK015948
LOCUS AK015948 1054 bp mRNA HTC 05-JUL-2001
DEFINITION Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:4930542L20, full insert sequence.
ACCESSION AK015948
VERSION AK015948.1 GI:12854490
KEYWORDS CAP trapper.
SOURCE Mus musculus (strain:C57BL/6J) adult male testis cDNA to mRNA.
clone_lib:RIKEN full-length enriched mouse cDNA library clone:4930542L20.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1054)
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Methods Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2 (bases 1 to 1054)
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE 3 (bases 1 to 1054)
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kochiwa,H., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsumura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer

```
JOURNAL   Genome research. 10 (11), 1757-1771 (2000)
 MEDLINE  20530913
  PUBMED  11076861
REFERENCE 4  (bases 1 to 1054)
  AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the
          FANTOM Consortium.
  TITLE   Functional annotation of a full-length mouse cDNA collection
  JOURNAL Nature 409, 685-690 (2001)
REFERENCE 5  (bases 1 to 1054)
  AUTHORS Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A.,
          Arakawa,T., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M.,
          Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F.,
          Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kato,H., Kawai,J.,
          Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T.,
          Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y.,
          Okido,T., Owa,C., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H.,
          Sasaki,D., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T.,
          Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
          Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A.,
          Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
  TITLE   Direct Submission
  JOURNAL Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
          Physical and Chemical Research (RIKEN), Laboratory for Genome
          Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
          RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
          Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
          URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
          Fax:81-45-503-9216)
COMMENT   Please visit our web site (http://genome.gsc.riken.go.jp/) for
          further details.
          cDNA library was prepared and sequenced in Mouse Genome
          Encyclopedia Project of Genome Exploration Research Group in Riken
          Genomic Sciences Center and Genome Science Laboratory in RIKEN.
          Division of Experimental Animal Research in Riken contributed to
          prepare mouse tissues. First strand cDNA was primed with a primer
          [5' GAGAGAGAGAACGATCCAACAGCTCTTTTTTTTTTTTTTTTTVN 3'], cDNA was
          prepared by using trehalose thermo-activated reverse transcriptase
          and subsequently enriched for full-length by cap-trapper. cDNA went
          through one round of normalization to Rot = 10.0 and subtraction to
          Rot = 100.0. Second strand cDNA was prepared with the primer
          adapter of sequence [5'
          GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCC 3'], cDNA was cleaved
          with BamHI and XhoI. Vector: a modified pBluescript KS(+) after
          bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3'
          end: BamHI. Host: DH10B.
FEATURES             Location/Qualifiers
     source          1..1054
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /db_xref="MGD:MGI:1908511"
                     /db_xref="MGD:MGI:1922446"
                     /clone="4930532L20"
                     /sex="male"
                     /tissue_type="testis"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="adult"
     CDS             110..949
                     /note="putative"
                     /codon_start=1
                     /protein_id="BAB30047.1"
                     /db_xref="GI:12854491"
                     /translation="MKFFFFPRGKRKIDDSHSHSSFVPISLAKIAPPNLSIGGCYPIP
                     DKHLKKLHKAATIGNEQKLKDYLEKKKYNVNGPDKRSFTPLHLACANGYINIVSLLIE
                     NQCKINVQDSENRTPLIKAVECQQESCATVLLLHGADPNLVDVYSNTALHYAVCGQNI
                     SLANKLLQYKANLEAKNKDGHTPLLLAVAENNENMVKFLLKKGADVNASDKNHRTAIM
                     IALIVEPTSSVKLLLQQDIDLAHKDIYGFTAFFYASFNCFTMYHHITANNENKKIEQ
                     TAY"
     polyA_signal    1033..1038
                     /note="putative"
     polyA_site      1054
                     /note="putative"
BASE COUNT      346 a    209 c    207 g    292 t      2 others
ORIGIN

  Query Match             18.8%;  Score 217.6;  DB 12;  Length 1054;
  Best Local Similarity   56.6%;  Pred. No. 3.1e 42;
  Matches  403;  Conservative    0;  Mismatches  309;  Indels    0;  Gaps    0;

Qy     382 ttcatggagcccaggtatcacgtccgtgggagaagatctggatcagatcacagatctgac 441
Db     215 TCCATAGGTGGAGGATATCCCCTTCGAGATAAGCACTTAAAGAAACTTCATAAAGCTGCT 274

Qy     442 tgatgggtaaagtcccagaaaggatctcatcgtcatgctcagggacactgacgtgaac 501
Db     275 ACCATTGCCAATGAGCAGAAGCTCAAGCATTACCTTGAACCCAAGAAATACAATGTGAAT 334

Qy     502 aagaaggacaagcaaaagaggatgctctaatcttggcttgcaatggaattcagaa 561
Db     335 GGCCCGGACAAAAGAAGCACAACACCTTTGCATTTGCCCTGTCCTAATGGATATACAAAT 394

Qy     562 gtagtaaaactcctgctggacagacgatgtcaacttaatgtccttgacaacaaaagagg 621
Db     395 ATTGTCTCTCTCTTAATTGAGAATCAATGCAAAATTAATGTCCAGGATAGTGAAAACAGA 454

Qy     622 acagctctgataaaggccgtacaatgccaggaagatgaatgtgcgttaatgttgctggaa 681
Db     455 ACCCCATTCATTAAGCCACTACACTGTCAACACCACAGTTGTCCTACGGTTCTTCTTCTT 514

Qy     682 catggcactgatccaaatattccagatgagtatggaaataccactctgcactacgctatc 741
Db     515 CATGGTGCAGATCCAAATCTGGTAGATGTTTATACTAATACGGCTCTCCATTATGCTGTT 574

Qy     742 tataatgaagataaattaatggccaaagcactgctcttatatggtgctgatatcgaatca 801
Db     575 TGTGGCCAAAATATTTCATTAGCCAATAAATTGCTTCAATATAAAGCCAATCTTGAAGCC 634

Qy     802 aaaaacaagcatggcctcacaccactgttacttggtgtacatgagcaaaacaccaagtc 861
Db     635 AAAAATAAGGACGGTCACACTCCGCTTTTACTTGCTGTTGCTGAAAATAATGAAAATATG 694

Qy     862 gtgaaattttaatcaagaaaaagcgaatttaaatgcactggatagatatggaaggact 921
Db     695 GTAAAATTCCTTCTGAACAAACCAGCAGATCTAAATGCATCACATAAAAACCACACAACA 754

Qy     922 gctctcatacttgctgtatgttgtggatcagcaagtatagtcagccttctacttgagcaa 981
Db     755 GCCATCATGATTGCTCTGATTGTTGAACCAACAAGTTCTGTAAAACTTCTTCTTCAGCAA 814

Qy     982 aatattgatgtatcttctcaagatctatctggacagacggccagagagtatgctgtttct 1041
Db     815 GATACGGACCTAGCCCACAAAGATATTTATGGATTTACAGCTGAGGAATATGCTTCATTT 874

Qy    1042 agtcatcatcatgtaatttgccagttactttctgactacaaagaaaaacaga 1093
Db     875 AATGGCTTTACTATGTATCACCATATAACTGCCAATAATGAGAACAAGAAGA 926


RESULT  8
BF675057
LOCUS       BF675057     822 bp    mRNA            EST       21-DEC-2000
DEFINITION  602136651F1 NIH_MGC_83 Homo sapiens cDNA clone IMAGE:4273141 5',
            mRNA sequence.
ACCESSION   BF675057
VERSION     BF675057.1  GI:11948952
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 822)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
```

```
            Email: cgapbs@mail.nih.gov
            Tissue Procurement: CLONETECH Laboratories, Inc.
             cDNA Library Preparation: CLONETECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1088  row: e  column: 14
            High quality sequence stop: 628.
FEATURES             Location/Qualifiers
     source          1..822
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4273141"
                     /clone_lib="NIH_MGC_83"
                     /lab_host="[illegible]"
                     /note="Organ: prostate; Vector: pDNR-LIB (Clontech);
                     Site 1: SfiI (ggccgcctcggcc); Site 2: SfiI (ggccattatggcc
                     ); 5' and 3' adaptors were used in cloning as follows: 5'
                     adaptor sequence:  5' CACGGCCATTATGGCC 3' and 3' adaptor
                     sequence:  5' ATTCTAGAGGCCGAGGCGGCCGACATG dT(30)BN 3'
                     (where B = A, C, or G and N = A, C, G, or T).  Average
                     insert size 1.4 kb (range 0.5-4.0 kb).  14/15 colonies
                     contained inserts  by PCR. This library was enriched for
                     full length clones and was constructed by Clontech
                     Laboratories (Palo Alto, CA)."
BASE COUNT      225 a    194 c    211 g    192 t
ORIGIN

  Query Match             18.5%;  Score 213.6;  DB 11;  Length 822;
  Best Local Similarity   64.4%;  Pred. No. 2.7e-41;
  Matches  366;  Conservative    0;  Mismatches  199;  Indels    4;  Gaps    3;

Qy        523 actgctctacatctggcctctgccaatggaattcagaagtagtaaaactcctgctggac 582
Db        102 ACTGCTCTACACTGGCTGTGCCAGTGCCATGTGAACTGGTCACTCTCCTGGTTAAC 161

Qy        583 agcagatgtcaacttaatgtccttgacaacaaaaagaggacagctctgataaaggccgta 642
Db        162 AGAAAATGCCAGATTGATATCTGTGACAAAGAAAACAGAACGCCTTTGATACAGGCTATC 221

Qy        643 [illegible] 702
Db        222 CATTGCCAGGAAGAGGCTTGTGCCGTTATTCTGCTGAAACATGGCGCCAATCCAAACCTT 281

Qy        703 [illegible] 762
Db        282 AAGGATATCTACGGCAACACTGCTCTCCATTATGCTGTGTATAGTGAGAGCACCTCACTG 341

Qy        763 [illegible] 822
Db        342 GCAGAAAAACTGCTTTCCCATGGTGCACACATTGAAGCACTGGACAAGGACAGTAATACC 401

Qy        823 [illegible] 881
Db        402 CCACTTTTATTCGCTATAATTTGCAAGAAAGAGAAACATGGTGGAATTTTTATTGAAAAA 461

Qy        882 [illegible] 941
Db        462 GAAAGCAAGTACACATGCCGTTGATAGGCTGACGGTCAGCTCTCATGCTTGCTGTATA 521

Qy        942 [illegible] 1000
Db        522 CTATGACTCACCAGGTAT GTCAGTATCCTTCTTACGCAAAATATTGATGTCTCCGGCTC 580

Qy       1001 [illegible] 1060
Db        581 AAGACATGTGTGGACGAGATGAGAAGATACGCTATTTTCATTCATTGACCACAATT 640

Qy       1061 [illegible] 1088
Db        641 CACACCACATTGGACCTAAACCGAAGA 668


RESULT 9
BG703856
LOCUS       BG703856               810 bp    mRNA            EST       07-MAY-2001
DEFINITION  602686889F1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:4819127 5',
            mRNA sequence.
ACCESSION   BG703856
VERSION     BG703856.1  GI:13976606
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 810)
  AUTHORS   NIH-MGC [illegible]
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
             cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10722  row: b  column: 24
            High quality sequence stop: 745.
FEATURES             Location/Qualifiers
     source          1..810
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4819127"
                     /clone_lib="NIH_MGC_95"
                     /tissue_type="hippocampus"
                     /lab_host="DH10B"
                     /note="Organ: brain; Vector: pBluescriptR (modified
                     pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (attead
                     ); oligo dT primed using primer 5' [illegible]VN 3',
                     size selected for  average insert size 2.5 kb and
                     normalized to Rot 5. This is a primary library enriched
                     for  full length clones and constructed using the
                     Cap trapper method (Carninci, in preparation).  Library
                     constructed by M. Brownstein (NIMH/NHGRI, National
                     Institutes of Health).  Note: this is a NIH_MGC Library."
BASE COUNT      253 a    151 c    191 g    215 t
ORIGIN

  Query Match             18.3%;  Score 211.6;  DB 11;  Length 810;
  Best Local Similarity   59.5%;  Pred. No. 8.3e-41;
  Matches  358;  Conservative    0;  Mismatches  244;  Indels    0;  Gaps    0;

Qy        395 ggtaccacgtccgtggagaaatctggacaagctccacagagctgcctggtgggtaaat 454
Db        146 GCTACAACCTTCGAGAAAAGGATTTAAAGAAACTTCACAGAGCTGCTTCAGTCGGGCATT 205

Qy        455 tccccagaaaggatctcatcgtcatgctcagggacactgacgtgaacaaaaggacaagc 514
Db        206 TGAAGAAGCTGAAGGAATACCTTCAGATCAAGAAATATGATTAAATATGCAGGACAAAA 265

Qy        515 aaaagaggactgctctacatctggcctctgccaatggaattcagaagtagtaaaactcc 574
Db        266 AATACAGAACACCTTTGCACCTAGCCTGTGCTAATGGACATACAGA[illegible]GTACTTTT[illegible] 325

Qy        575 tgctggacagacgatatgtcaacttaatgtccttgacaacaaaaa[illegible]actctgataa 634
Db        326 TAATTGAGCAACAATGCAAAATAAATGTCCGGGATAGTGAAAACAAATGCCATTGATTA 385

Qy        635 aggccgtacaatgccaggaagatgaatgtgcgttaatgttgctggaaacatggcactgatc 694
```

```
Db     386 AGGCAGTACAGTGTCAAAATGAGGATGTGCTACTATTCTTCTAAACTTTGGTGCAGACC 445

Qy     695 caaatattccagatgagtatggaaataccactctgcactacgctatctataatgaagata 754

Db     446 CACATCTGAGGGATATTCGTTATAATACTGTTCTTCACTATGCTGTTTGTGGTCAAAGTT 505

Qy     755 aattaatggccaaagcactgctcttatatggtgctgatatcgaatcaaaaaacaagcatg 814

Db     506 TGTCATTAGTTGAAAAACTGCTTGAATACGAAGCTGATCTTGAAGCGAAAAATAAGGATG 565

Qy     815 gcctcacaccactgttacttggtgtacatgagcaaaaacagcaagtcgtgaaattttaa 874

Db     566 GGTATACTCCACTATTAGTTGCCGTTATTAACAATAATCCAAAAATGGTAAAATTTCTTC 625

Qy     875 tcaagaaaaaagcgaatttaaatgcactggatagatatggaaggactgctctcatacttg 934

Db     626 TGGAGAAAGGGGCTGATGTGAATGCTTCAGATAATTATCAAAGAACAGCCCTTATTCTTG 685

Qy     935 ctgtatgttgtggatcagcaagtatagtcagccttctacttgagcaaaatattgatgtat 994

Db     686 CTGTCAGTGGTGAACCACCATGTTTAGTAAAGCTTCTTCTTCAGCAACGTGTGGAATTAT 745

Qy     995 ct 996

Db     746 GT 747
```

RESULT 10
BG721483
LOCUS   BG721483   907 bp   mRNA   EST   08-MAY-2001
DEFINITION 602695061F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4827350 5',
            mRNA sequence.
ACCESSION   BG721483
VERSION   BG721483.1 GI:14000670
KEYWORDS   EST.
SOURCE   human.
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 907)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
             cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10743 row: i column: 15
            High quality sequence stop: 758.
FEATURES   Location/Qualifiers
     source   1..907
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4827350"
                     /clone_lib="NIH_MGC_97"
                     /lab_host="DH10B"
                     /note="Organ: testis; Vector: pBluescriptR (modified
                     pBluescript KS+); Site_1: BamHI; Site_2: SalI XhoI (gtcgag
                     ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTTVN-3',
                     size-selected for average insert size 2.2 kb and
                     normalized to ROT 5. This is a primary library enriched
                     for full-length clones and constructed using the
                     Cap-trapper method (Carninci, in preparation). Library
                     constructed by M. Brownstein (NIMH/NHGRI, National
                     Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT   287 a   166 c   224 g   229 t   1 others
ORIGIN

  Query Match   18.1%; Score 209.6; DB 11; Length 907;
  Best Local Similarity   59.3%; Pred. No. 2.6e-40;
  Matches 356; Conservative   0; Mismatches 244; Indels   0; Gaps   0;

```
Qy     395 ggtaccacgtccgtggagaagatctggacaagctccacagagctgcctggtggggtaaag 454

Db     167 CCTACAACCTTCGAGAAAAGCATTTAAAGAAACTTCACAGAGCTGCTTCAGTCGGGGATT 226

Qy     455 tccccaggaaggatctcatcgtcatgctcggggacactgacgtggacaagaacgacaagc 514

Db     227 TCAAGAACCTGAAGGAATACCTTCAGATCAAGAAATATGATGTAAATATGCACGACAAAA 286

Qy     515 aaaagatgactgctctacatctggcctctgccaatgggaattcagaagtagtaaaactcc 574

Db     287 AATACAGAACACCTTTGCACCTAGCCTGTGCTAATGGACATACAGATGTTGTACTTTTCC 346

Qy     575 tgctggacagacgatgtcaacttaatgtccttgacaacaaaaagaggacagctctgataa 634

Db     347 TAATTGAGCAACAATGCAAAATAAATGTCCGGGATAGTGAAAACAAATCCCCATTGATTA 406

Qy     635 aggcggtacaatgcaggaagatgaatgtgccttaatgttgctgagacatggcactgatc 694

Db     407 AGGCAGTACAGTGTCAAAATGAGGATTGTGCTACTATTCTTCTAAACTTTGGTGCAGACC 466

Qy     695 caaatattccagatgagtatggaaataccactctgcactacgctatctataatgaagata 754

Db     467 CAGATCTGAGGGATATTCGTTATAATACTGTTCTTCACTATGCTGTTTGTGGTCAAAGTT 526

Qy     755 aattaatggccaaagcactgctcttatatggtgctgatatcgaatcaaaaaacaagcatg 814

Db     527 TGTCATTAGTTGAAAAACTGCTTGAATACGAAGCTGATCTTGAAGCGAAAAATAAGGATG 586

Qy     815 gcctcacaccactgttacttggtgtacatgagcaaaaacagcaagtcgtgaaattttaa 874

Db     587 GGTATACTCCACTATTAGTTGCCGTTATTAACAATAATCCAAAAATGGTAAAATTTCTTC 646

Qy     875 tcaagaaaaaagcgaatttaaatgcactggatagatatggaaggactgctctcatacttg 934

Db     647 TGGAGAAAGGGGCTGATGTGAATGCTTCAGATAATTATCAAAGAACAGCCCTTATTCTTG 706

Qy     935 ctgtatgttgtggatcagcaagtatagtcagccttctacttgagcaaaatattgatgtat 994

Db     707 CTGTCAGTGGTGAACCACCATGTTTAGTAAAGCTTCTTCTTCAGCAAGGTGTGGACTTAT 766
```

RESULT 11
BG716974
LOCUS   BG716974   916 bp   mRNA   EST   08-MAY-2001
DEFINITION 602689005F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4821576 5',
            mRNA sequence.
ACCESSION   BG716974
VERSION   BG716974.1 GI:13996161
KEYWORDS   EST.
SOURCE   human.
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 916)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
             cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:

```
            http://image.llnl.gov
            Plate: LLAM10728  row: i  column: 01
            High quality sequence stop: 832.
FEATURES             Location/Qualifiers
     source          1..916
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4823376"
                     /clone_lib="NIH_MGC_97"
                     /lab_host="DH10B"
                     /note="Organ: testis; Vector: pBluescriptR (modified
                     pBluescript KS+); Site 1: BamHI; Site 2: SalI XhoI (gtcgac
                     ), oligo dT primed using primer 5' TTTTTTTTTTTTTTTTVN-3',
                     size selected for  average insert size 2.2 kb and
                     normalized to ROT 5. This is a primary library enriched
                     for full length clones and constructed using the
                     Cap trapper method (Carninci, in preparation).  Library
                     constructed by M. Brownstein (NIMH/NHGRI, National
                     Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT      282 a    174 c    218 g    242 t
ORIGIN

  Query Match             18.0%;  Score 207.8;  DB 11;  Length 916;
  Best Local Similarity   59.0%;  Pred. No. 7e-40;
  Matches  ...;  Conservative    0;  Mismatches  247;  Indels    0;  Gaps    0;

Qy        387 ggagccaggtaccacgtccgtggagaagatctggacaagctccacagagctgcctggtg 446
Db        160 TTAGTTAGGCTACAACCTTGAGAAAAGGATTTAAAGAAACTTCACAGAGCTGCTTCAGT 219

Qy        447 ggatttaagtccccggaaggatctcatcgtcatgcttcaggggacactgacgtggacaagaa 506
Db        220 CGGGGATTTGAAGAAGCTGAAGGAATACCTTCAGATCAAGAAATATGATGTAAATATGCA 279

Qy        507 ggacaagcaaaagaggactgctctacatctggcctctgccaatgggaattcagaagtagt 566
Db        280 TGACAAAAAATACAGAACACCTTTGCACCTAGCCTGTGCTAATGGACATACAGATGTTGT 339

Qy        567 aagactcctgctggacagacgatgtcaagcttaatgtccttgacaacaaaaaggaggacagc 626
Db        340 ACTTTTCCTAATTGAGCAACAATGCAAAATAAATGTCCGGGATAGTGAAAACAAATCCCC 399

Qy        627 tctgataaaggccgtacaatgccaggaagatgaatgtgcgttaatgttgctggaacatgg 686
Db        400 ATTGATTAAGGCAGTACAGTGTCAAAATGAGGATTGTGCTACTATTCTTCTAAACTTTGG 459

Qy        687 cactgatccaaatattccagatgagtatggaaataccactctgcactacgctatctataa 746
Db        460 TTCAGACCCAGATCTGAGGATATTCGTTATAATACTGTTCTTCACTATGCTGTTTGTGG 519

Qy        747 tggagataaattaatggccaaagcactgctcttatatggtgctgatatcgaatcaaaaaa 806
Db        520 TCAAAGTTTTGTCATTAGTTGAAAAACTGCTTGAATACGAAGCTGATCTTGAAGCGAAAAA 579

Qy        807 caagcatggcctcacaccactgttacttggtgtacatgagcaaaaacagcaagtcgtgaa 866
Db        580 TAATGATGGGTATACTCCACTATTAGTTGCCGTTATTAACAATAATCCAAAAATGGTAAA 639

Qy        867 atttttaatcaagaaaaaagcgaatttaaatgcactggatagatatggaaggactgctct 926
Db        640 ATTTCTTCTGGAGAAAGGGGCTGATGTGAATGCTTCAGATAATTATCAAAGAACAGCCCT 699

Qy        927 catacttgctgtatgttgtggatcagcaagtatagtcagccttctacttgagcaaaatat 986
Db        700 TATTCTTGCTGTTCAGTGGTGAACCACCACATTTAGTAAAGCTTCTTCTTCAGCAAGTGT 759

Qy        987 tga 989
Db        760 TGA 762


RESULT 12
AK016475
LOCUS       AK016475     1832 bp    mRNA            HTC       05-JUL-2001
DEFINITION  Mus musculus adult male testis cDNA, RIKEN full-length enriched
            library, clone:4931423N10, full insert sequence.
ACCESSION   AK016475
VERSION     AK016475.1  GI:12855225
KEYWORDS    CAP trapper.
SOURCE      Mus musculus (strain C57BL/6J) adult male testis cDNA to mRNA,
            clone_lib:RIKEN full-length enriched mouse cDNA library
            clone:4931423N10.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1832)
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE     High-efficiency full-length cDNA cloning
  JOURNAL   Methods in enzymology. 303, 19-44 (1999)
  MEDLINE   99279253
   PUBMED   10349636
REFERENCE   2  (bases 1 to 1832)
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL   Genome research. 10 (10), 1617-1630 (2000)
  MEDLINE   20499374
   PUBMED   11042159
REFERENCE   3  (bases 1 to 1832)
  AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
            Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
            Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
            Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
            Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
            Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
            Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
  TITLE     RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer
  JOURNAL   Genome research. 10 (11), 1757-1771 (2000)
  MEDLINE   20530913
   PUBMED   11076861
REFERENCE   4  (bases 1 to 1832)
  AUTHORS   The RIKEN Genome Exploration Research Group Phase II Team and the
            FANTOM Consortium.
  TITLE     Functional annotation of a full-length mouse cDNA collection
  JOURNAL   Nature 409, 685-690 (2001)
REFERENCE   5  (bases 1 to 1832)
  AUTHORS   Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Arai,A.,
            Arakawa,T., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M.,
            Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F.,
            Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kato,H., Kawai,J.,
            Kojima,Y., Konno,H., Kondo,M., Koya,S., Kurihara,C., Matsuyama,T.,
            Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y.,
            Okido,T., Owa,C., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H.,
            Sasaki,D., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T.,
            Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
            Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A.,
            Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
            Physical and Chemical Research (RIKEN), Laboratory for Genome
            Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
            RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
            Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
            Fax:81-45-503-9216)
COMMENT     Please visit our web site (http://genome.gsc.riken.go.jp/) for
            further details.
            cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues. First strand cDNA was primed with a primer
            [5'-GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTVN-3'], cDNA was
```

```
            prepared by using trehalose thermo-activated reverse transcriptase
            and subsequently enriched for full-length by cap-trapper. cDNA went
            through one round of normalization to Rot = 10.0 and subtraction to
            Rot = 100.0. Second strand cDNA was prepared with the primer
            adapter of sequence [5'
            GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCC 3']. cDNA was cleaved
            with BamHI and XhoI. cDNA of size comprised longer than 7 kb was
            selected before cloning. Vector. a modified pBluescript KS(+) after
            bulk excision from Lambda FLC 1. Cloning sites, 5' end: SalI; 3'
            end: BamHI. Host: DH10B.
FEATURES             Location/Qualifiers
     source          1..1832
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /db_xref="MGD:MGI:1903457"
                     /db_xref="MGD:MGI:1918231"
                     /clone="4931423N10"
                     /sex="male"
                     /tissue_type="testis"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="adult"
     CDS             338..1486
                     /note="putative"
                     /codon_start=1
                     /protein_id="BAB30257.1"
                     /db_xref="GI:12855226"
                     /translation="MGGKYSRGFERRPLCFPGFSALCCCQKTNFEETPLGFCDIKARR
                     SNLSFGREDPAYLDMAYFPDNDLHMAACAGDLPFVRLYFTLGKYEVNHRDRENRNAMH
                     FACFYGHLELVIYLWRRGCEINVCDNHNITPLMKAVQSWEDKIVCFLLEHHANLHIKD
                     SMGNTALHYAVYSGNLATAARLLQYGADIEERTKDNLTPLLLALRFNPLKMAQFLVRM
                     EASVHAVDSQRRNSLMYAVRCDSPVMVNLLLQQGVDINLKDLFCWTALRYAIEGDRDV
                     RTMLLELRKRNRAFQSSENSSIRLINKADAEVTSPTANEKVIEMQDPFTTNEASVEKT
                     QTRRPAESFLEVFSLEEEISYKETINNLQQNPCQRSAKIDAATENRSP"
     polyA_signal    1808..1813
                     /note="putative"
     polyA_site      1832
                     /note="putative"
BASE COUNT      574 a    336 c    448 g    474 t
ORIGIN

  Query Match             18.0%;  Score 207.4;  DB 12;  Length 1832;
  Best Local Similarity   58.0%;  Pred. No. 9.8e-40;
  Matches  367;  Conservative    0;  Mismatches  266;  Indels    0;  Gaps    0;

Qy        425 agctccacagagctgcctggtggggtaaagtcccagaaaggatctcatcgtcatgctca 484
              | || ||||  || |||||     ||   |   | ||        |   ||   | | |  ||
Db        534 ATCTGCACATGGCGGCCTGTGCAGGCGATTTACCTTTTGTGCGGCTCTACTTTACTCTGG 593

Qy        485 gggacactgacgtgaccagaaggacaagcaaaagaggctgctctacatctggcctctg 544
              |  |     || || ||   | |   |||| | | || |||| |||   | ||   | ||||
Db        594 GTAAATACGAAGTCAATCACAGAGACAGGCAGAATAGGAATGCCATGCACTTTGCCTGCT 653

Qy        545 ccaatgggaattcagaagtagtaaaactcctgctggacagacgatgtcaacttaatgtcc 604
               | ||||   |||  |||| | | || |  |   ||    |      |  | | ||| |   |||||||
Db        654 TCTATGGACATTTAGAATTGGTGATATATCTCTGGAGGCGTGGTTGTGAGATTAATGTGT 713

Qy        605 ttgacaacaaaaagaggacagctctgataaaggccgtgcaatgccaggaagatgaatgtg 664
               ||||||   |  ||  |    |||  | ||||| ||||| || ||| ||   ||||||| ||     |
Db        714 CTCACAATCATAACATCACACCCCTCATCAAGCCTGTCCAAACCTGGGAAGATAAAATCG 773

Qy        665 cgttaatgttgctggaacatggcactgatccaaatattccagatgagtatggaaatacca 724
               ||    |    |||| || ||          || | |     | |||  |||          || || ||
Db        774 TGTGTTTTCTGCTAGAGCACCATGCTAACCTGCACATTAAAGACAGCATGGGTAACACTG 833

Qy        725 ctctgcactacgctatctataatgaagataaattaatggccaaagcactgctcttatatg 784
              |||||||||| ||| ||||||     |     ||    |    || |||          |||||     || |
Db        834 CTCTGCACTATGCTGTCTATAGCGGGAATTTAGCAACGGCTGCCAGGCTGCTACAGTACG 893

Qy        785 gtgctgatatcgaatcaaaaaacaagcatggcctcacaccactgttacttggttacatg 844
              | || ||||| ||     ||   |    ||    |       ||||||||||| || |||| | |      |
Db        894 GAGCAGATATTGAGGAAAGGACAAAAGACAACCTCACACCACTCTTGCTTGCTCTCAGGG 953

Qy        845 agcaaaaacagcaagtcgtgaaattttaatcaagaaaaagcgaatttaaatgcactgg 904
               || | | || | |   | |    |||| ||| | | ||      |||| |     |   | ||| | |
Db        954 AGAACAGACTGAAGATGGCACAATTCTTAGTAAGGATGGAAGCAAGCGTGCATGCAGTTG 1013

Qy        905 atagatatggaaggactgctctcatacttgctgtatgttgtggatcagcaagtatagtca 964
               |    | |   |||| | | |||||||    ||||||| | |||| ||| ||   || ||||
Db       1014 ACTCACAGAGAAGAAATTCTCTCATGTATGCTGTAAGGTGTGACTCACCAGTTATGGTCA 1073

Qy        965 gccttctacttgagcaaaatattgatgtatcttctcaagatctatctggacagacggcca 1024
               |||| | ||| | |||   | | ||  ||  ||    ||||| ||| |||| ||| ||
Db       1074 ACCTTATTCTTCAACAAGGTGTCGACATAAATTTGAAAGATTTATTTGGATGGACTGCTT 1133

Qy       1025 gagagtatgctgtttctagtcatcatcatgtaa 1057
               |    |||||| ||    || ||| | |||| |
Db       1134 TACGTTATGCTATTGAAGGTGATCGTGATGTGA 1166


RESULT 13
AK016868
LOCUS       AK016868     1203 bp    mRNA            HTC       05-JUL-2001
DEFINITION  Mus musculus adult male testis cDNA, RIKEN full-length enriched
            library, clone:4933422A18, full insert sequence.
ACCESSION   AK016868
VERSION     AK016868.1  GI:12855833
KEYWORDS    CAP trapper.
SOURCE      Mus musculus (strain:C57BL/6J) adult male testis cDNA to mRNA,
            clone_lib:RIKEN full-length enriched mouse cDNA library
            clone:4933422A18.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1203)
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE     High-efficiency full-length cDNA cloning
  JOURNAL   Methods in enzymology. 303, 19-44 (1999)
  MEDLINE   99279253
   PUBMED   10349636
REFERENCE   2  (bases 1 to 1203)
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL   Genome research. 10 (10), 1617-1630 (2000)
  MEDLINE   20499374
   PUBMED   11042159
REFERENCE   3  (bases 1 to 1203)
  AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
            Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
            Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
            Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
            Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
            Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
            Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
  TITLE     RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer
  JOURNAL   Genome research. 10 (11), 1757-1771 (2000)
  MEDLINE   20530913
   PUBMED   11076861
REFERENCE   4  (bases 1 to 1203)
  AUTHORS   The RIKEN Genome Exploration Research Group Phase II Team and the
            FANTOM Consortium.
  TITLE     Functional annotation of a full-length mouse cDNA collection
  JOURNAL   Nature 409, 685-690 (2001)
REFERENCE   5  (bases 1 to 1203)
  AUTHORS   Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A.,
            Arakawa,T., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M.,
            Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F.,
            Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kato,H., Kawai,J.,
            Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T.,
            Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y.,
            Okido,T., Owa,C., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H.,
            Sasaki,D., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T.,
```

```
            Osato,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
            Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A.,
            Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
            Physical and Chemical Research (RIKEN), Laboratory for Genome
            Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
            RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
            Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
            URL:http://[illegible], Tel:81-45-503-9222,
            Fax:81-45-503-9216)
COMMENT     Please visit our web site ([illegible]) for
            further details.
            cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues. First strand cDNA was primed with a primer
            [5'-GAGACAGACAAGGATCCAAGAG[illegible]TTTTTTTTTTTTTTVN-3'], cDNA was
            prepared by using trehalose thermo-activated reverse transcriptase
            and subsequently enriched for full-length by cap-trapper. cDNA went
            through one round of normalization to Rot = 10.0 and subtraction to
            Rot = 100.0. Second strand cDNA was prepared with the primer
            adapter of sequence [5'-
            GAGAGAGAGATCTCGAGTTAATTAAATTAATCCCCCCCCCCCC-3']. cDNA was cleaved
            with BamHI and XhoI. cDNA of size comprised between 0.5 and 3 kb
            was selected before cloning. Vector: a modified pBluescript KS(+)
            after bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI;
            3' end: BamHI. Host: DH10B.
FEATURES             Location/Qualifiers
     source          1..1203
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /db_xref="MGD:MGI:1894656"
                     /db_xref="MGD:MGI:1914825"
                     /clone="4933422A18"
                     /sex="male"
                     /tissue_type="testis"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="adult"
     CDS             192..1203
                     /note="putative"
                     /codon_start=1
                     /protein_id="BAB30474.1"
                     /db_xref="[illegible]"
                     /translation="MEE[illegible]VE[illegible]VEEEEEEE[illegible]DVEEEN[illegible]LLEEMEA[illegible]SVENE[illegible]EYE
                     PLRKIHQAASHG[illegible]EKL[illegible]KM[illegible]KHSVHDRD[illegible]KLRIALHIACACGQVEVVSLLLRNN
                     CDIDAADMNEIIPLMEAV[illegible]NW[illegible]Y[illegible]V[illegible]TT[illegible]K[illegible]ADENEE[illegible]KNCNTSLHYAVSEDNQTL
                     AE[illegible]L[illegible]KYSANME[illegible]NE[illegible]TPLLIALEENE[illegible]EMAKE[illegible]VEMGANI[illegible]VE[illegible]MEENTLLYA
                     TRWDSKEMVNLLL[illegible]VE[illegible]WIALRYATE[illegible]RE[illegible]MNYILMLEKKNKE
                     GIPEYKEEEDKSLDKHPNNE[illegible]SNENA[illegible]KE[illegible]LL[illegible]SKYVLSTCPNNIHGAC
                     EES"
BASE COUNT      [illegible] a    208 c    276 g    320 t
ORIGIN

  Query Match             17.9%;  Score 206.2;  DB 12;  Length 1203;
  Best Local Similarity   56.6%;  Pred. No. 1.8e-39;
  Matches  382;  Conservative    0;  Mismatches  293;  Indels    0;  Gaps    0;

Qy        413 [illegible]ggacaagctccacagagctgcctggtggggtaaagtcccagaaaaggatctca 472
Db        322 AACCTCTCAGAAAGATCCACCAGGCTGCAAGCCACGGTGATACAGAGACTACAGAGAA 381

Qy        473 tcgtcatgctcag[illegible]acactgacgtgaacaaaggacaagcaaaagagaactgctctac 532
Db        382 TGATAAGTCTTGGAAACACAGTGTGCATGACAGAGACTTTAAGGAAAGGACTGCCCTGC 441

Qy        533 at[illegible]gcctct[illegible]caatgggatttcagaagtagt[illegible]aaactcctgctggacagacgatgtc 592
Db        442 ATTTGCCTGTGCCTGTGGCCAGGTGGAAGTAGTAAGTATTCTATTGAGAAATAACTGTG 501

Qy        593 aacttaatgtccttgacaacaaaaagaggacagctctgataaag[illegible] 652
Db        502 ATATTGACGCTGCTGACATGAACTTCATCACGCCTCTGATGAAGGCTGTTCAAAACTGGA 561

Qy        653 aagatgaatgtgcgttaatgttgctgggaacatggcactgat[illegible] 712
Db        562 CATATGAATGTGTGTGCATATTATTGAAGCATGGTGCTGATCCAAACCGAAAGGACAAAA 621

Qy        713 atggaaataccactctgcactacgctatctataatgaagataatt[illegible] 772
Db        622 ATGGCAATACTTCTCTTCACTACGCTGTGTCAGAAGATAACCAGACACTCGCCAAATGCT 681

Qy        773 tgctcttatatggtgctgatatcgaatcaaaaaacaagcat[illegible] 832
Db        682 TGCTTAAGTACAGTGCAAACATGGAACAAAAGAACAAGGATGGCTTACACCACTTTAC 741

Qy        833 ttggtgtacatgagcaaaaacagcaagtcgtgaaatttttaat[illegible] 892
Db        742 TGGCTCTCAAGGAAAACAAAATAGAGATGGCAAAGTTTTAGTAAAGATGGGGCAAATA 801

Qy        893 taaatgcactggatagatatggaaggactgctctcatacttgct[illegible] 952
Db        802 TACATGTATTTGATGACATGAGAAGAAACACACTCTTATATGCCATAAGATGGACTCAA 861

Qy        953 caagtatagtcagccttctacttgagcaaaatattgatgtatcttctcaagatctatctg 1012
Db        862 AAGATATGGTCAACCTACTTTTGGATGAAGGCATTGACTTTTTCTTTAGAGATGTATTTG 921

Qy       1013 gacagacggccagagagtatgctgtttctagtcatcatcatgtaatttg[illegible] 1072
Db        922 GGTGGACTGCATTACGGTATGCTATTGAAGGCACAAGTAAAGGTAGTGGAGAGATTCTTA 981

Qy       1073 ctgactacaaaaaaa 1087
Db        982 TGAACTATGATGAAA 996


RESULT 14
AK015247
LOCUS       AK015247     1837 bp    mRNA            HTC       05-JUL-2001
DEFINITION  Mus musculus adult male testis cDNA, RIKEN full-length enriched
            library, clone:4930430A15, full insert sequence.
ACCESSION   AK015247
VERSION     AK015247.1  GI:12853508
KEYWORDS    CAP trapper.
SOURCE      Mus musculus (strain [illegible]) adult male [illegible] cDNA [illegible]
            clone_lib:RIKEN full-length enriched mouse cDNA library
            clone:4930430A15.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1837)
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE     High-efficiency full-length cDNA cloning
  JOURNAL   Methods in enzymology. 303, 19-44 (1999)
  MEDLINE   99279253
   PUBMED   10349636
REFERENCE   2  (bases 1 to 1837)
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL   Genome research. 10 (10), 1617-1630 (2000)
  MEDLINE   20499374
   PUBMED   11042159
REFERENCE   3  (bases 1 to 1837)
  AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
            Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
            Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
            Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
            Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
            Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
            Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
```

```
  TITLE       RIKEN integrated sequence analysis (RISA) system--384-format
              sequencing pipeline with 384 multicapillary sequencer
  JOURNAL     Genome research. 10 (11), 1757-1771 (2000)
  MEDLINE     20530913
   PUBMED     11076861
REFERENCE     4  (bases 1 to 1837)
  AUTHORS     The RIKEN Genome Exploration Research Group Phase II Team and the
              FANTOM Consortium.
  TITLE       Functional annotation of a full-length mouse cDNA collection
  JOURNAL     Nature 409, 685-690 (2001)
REFERENCE     5  (bases 1 to 1837)
  AUTHORS     Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Arai,A., Aono,H.,
              Arakawa,T., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M.,
              Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F.,
              Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kato,H., Kawai,J.,
              Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T.,
              Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y.,
              Okido,T., Owa,C., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H.,
              Sasaki,D., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T.,
              Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
              Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A.,
              Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
  TITLE       Direct Submission
  JOURNAL     Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
              Physical and Chemical Research (RIKEN), Laboratory for Genome
              Exploration Research Group, RIKEN Gemomic Sciences Center (GSC),
              RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
              Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
              URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
              Fax:81-45-503-9216)
COMMENT       Please visit our web site (http://genome.gsc.riken.go.jp/) for
              further details.
              cDNA library was prepared and sequenced in Mouse Genome
              Encyclopedia Project of Genome Exploration Research Group in Riken
              Genomic Sciences Center and Genome Science Laboratory in RIKEN.
              Division of Experimental Animal Research in Riken contributed to
              prepare mouse tissues. First strand cDNA was primed with a primer
              [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTVN 3'], cDNA was
              prepared by using trehalose thermo-activated reverse transcriptase
              and subsequently enriched for full-length by cap trapper. cDNA went
              through one round of normalization to Rot = 10.0 and subtraction to
              Rot = 100.0. Second strand cDNA was prepared with the primer
              adapter of sequence [5'
              GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCC 3']. cDNA was cleaved
              with BamHI and XhoI. Vector: a modified pBluescript KS(+) after
              bulk excision from Lambda FLC 1. Cloning sites, 5' end: SalI; 3'
              end: BamHI. Host: DH10B.
FEATURES               Location/Qualifiers
     source            1..1837
                       /organism="Mus musculus"
                       /strain="C57BL/6J"
                       /db_xref="taxon:10090"
                       /db_xref="MGD:MGI:1897149"
                       /db_xref="MGD:MGI:1914825"
                       /clone="4930430A15"
                       /sex="male"
                       /tissue_type="testis"
                       /clone_lib="RIKEN full-length enriched mouse cDNA library"
                       /dev_stage="adult"
     CDS               188..1645
                       /note="putative"
                       /codon_start=1
                       /protein_id="BAB29763.1"
                       /db_xref="GI:12853509"
                       /translation="MKKIIVHGVKFKIFLGFCIVPRKNLLDFLPDMKSISVPNEPFYK
                       PLRKIHQAASDGDIFRLQRMISLGKHSVHDPDFKERTALHFACACGQVEVVSILLRNN
                       CDIDAADMNFITPLMKAVQNWTYECVCILLKHGADPNRKDKNGNTSLHYAVSEDNQTL
                       AKCLLKYSANMFQKNKDGFTPLLLALKENKIEMAKFLVKMGANIHVFDDMERNTLLYA
                       IRWDSKDMVNLLLDEGIDFFFRDVFGWTALRYAIEGTSKGSREILMNYDEMLRRKNKD
                       GIPEYKFFEDKSLDKHPNNPISGCIFPKSNENACKFDDELDISKYVLGTCPNNIHGAG
                       EEGDKSETKKAADVQVWTNPTLKSSTKGKVPNLNEVCTMTESHHFIKESDLEIESLVE
                       IFSYDNEDKQSMSDDGHFSGTTPKWCQNDTNFKITVQDCMYPNLKSPTEGKIPTRSKA
                       GKMAESAHFIIESDLEIESLEETVSCDSEDKQLLV"
     polyA_signal      1817..1822
                       /note="putative"
     polyA_site        1837
                       /note="putative"
BASE COUNT      635 a    302 c    398 g    502 t
ORIGIN

  Query Match            17.9%;  Score 206.2;  DB 12;  Length 1837;
  Best Local Similarity  56.6%;  Pred. No. 1.9e-39;
  Matches  282;  Conservative    0;  Mismatches  293;  Indels    0;  Gaps    0;

Qy        413 aagatctggacaagctcacagagctgcctggtggggtaaagtccccagaaaggatctca 472
              || |||    ||| |||||    |||||  |     ||| |        |||    |    |
Db        318 AACCTCTCAGAAAGATCCACCAGGCTGCAAGCGACGGTGATACAGAGAGACTACAGAGAA 377

Qy        473 tcgttatgctcaggggacactgacgtgaacaaggaggacaagcaaaagaggactgctctac 532
              |  |  |  ||  || |         ||| |   | |  |||     |  | |||||||| || |
Db        378 TGATAAGTCTTGGGAAACACAGTGTGCATGACAGAGACTTTAAGGAAAGGACTGCCCTGC 437

Qy        533 atctggcctctgccaatgggattcagaagtagtaaaactcctgctggacagacgatgtc 592
              || | |||| ||||  |||   |      |||||||||     | ||  ||    |       |||
Db        438 ATTTTGCCTGTGCCTGTGGCCAGGTGGAAGTAGTAAGTATTCTATTGAGAAATAACTGTG 497

Qy        593 aacttaatgtccttgacaacaaaaagaggacagctctgataaaggccgtacaatgccagg 652
              |  || |  |     |||||  ||       |  ||  ||||||| ||||| || |||   |   |
Db        498 ATATTGACGCTGCTGACATGAACTTCATCACGCCTCTGATGAAGGCTGTTCAAAACTGGA 557

Qy        653 aagatgaatgtgcgttaatgttgctgaacatgcactgatccaaatattccaaatgaat 712
               | |||||||||| ||   || ||    || | |||||   |||||||||              | |
Db        558 CATATCAATGTGTGTGCATATTATTGAACCATGCTCCTGATCCAAACCGAAAGCACAAAA 617

Qy        713 atggaaatacccactctgcactacgtatctataatgaagataattaatggccaaagcac 772
              |||| |||||   |||| ||||||||| | |      | ||   |    |     | | ||| ||
Db        618 ATGGCAATACTTCTCTTCACTACGCTGTGTCAGAAGATAACCAGACACTCGCCAAATGCT 677

Qy        773 tgctcttatatggtgctgatatcgaatcaaaaaacaagcatggcctcacaccactgttac 832
              ||||      ||   ||||   |  ||  |||    ||| |||||| ||||| | |||||| | ||||
Db        678 TGCTTAAGTACAGTGCAAACATGGAACAAAAGAACAAGGATGGCTTTACACCACTTTTAC 737

Qy        833 ttggtgtacatgagcaaaaacagcaagtcgtgaatttttaatcaagaaaaacgaatt 892
              |  |  |  |     |  ||    |  |||      |   |   |   ||  |||||| | ||||        || |||
Db        738 TGGCTCTCAAGGAAAACAAAATAGAGATGGCAAAGTTTTTAGTAAAGATGGGGGCAAATA 797

Qy        893 taaatgcactggatagatatggaaggactgctctcatacttgctgtatgttgtggatcag 952
              || ||| | | |||          |||| |     | ||| ||   |||    || |  || |   |||
Db        798 TACATGTATTTGATGACATGAGAAGAAACACACTCTTATATGCCATAAGATGGCACTCAA 857

Qy        953 caagtatagtcagccttctacttgagcaaaatattgatgtatcttctcaagatctatctg 1012
               |   |||  ||||  |||  ||   |  ||   ||     |||||    |   |   |   ||||   || ||
Db        858 AAGATATGGTCAACCTACTTTTGGATGAAGGCATTGACTTTTTCTTTAGAGATCTATTTG 917

Qy       1013 gacagacggccagagagtatgctgtttctagtcatcatcatgtaatttgccagttacttt 1072
              |    |||  ||    |    ||||||| ||      |         |  |  |    |  |  |    ||  |  |||
Db        918 GGTGGACTGCATTACGGTATGCTATTGAAGGCACAACTAAAGGTAGTCGAGAGATTCTTA 977

Qy       1073 ctgactacaaagaaa 1087
                 ||||   |  ||||
Db        978 TGAACTATGATGAAA 992


RESULT 15
AK005925
LOCUS       AK005925     1209 bp    mRNA           HTC       05-JUL-2001
DEFINITION  Mus musculus adult male testis cDNA, RIKEN full-length enriched
            library, clone:1700012M14, full insert sequence.
ACCESSION   AK005925
VERSION     AK005925.1  GI:12838759
KEYWORDS    CAP trapper.
SOURCE      Mus musculus (strain:C57BL/6J) adult male testis cDNA to mRNA,
            clone_lib:RIKEN full-length enriched mouse cDNA library
            clone:1700012M14.
  ORGANISM  Mus musculus
```

```
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1209)
  AUTHORS   Carninci,P. and Hayashizaki,Y
  TITLE     High-efficiency full-length cDNA cloning
  JOURNAL   Methods in enzymology 303, 19-44 (1999)
  MEDLINE   [illegible]
   PUBMED   10349636
REFERENCE   2  (bases 1 to 1209)
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL   Genome research. 10 (10), 1617-1630 (2000)
  MEDLINE   [illegible]
   PUBMED   [illegible]
REFERENCE   3  (bases 1 to 1209)
  AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
            Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
            Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
            Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
            Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
            Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
            Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
  TITLE     RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer
  JOURNAL   Genome research. 10 (11), 1757-1771 (2000)
  MEDLINE   [illegible]
   PUBMED   11076861
REFERENCE   4  (bases 1 to 1209)
  AUTHORS   The RIKEN Genome Exploration Research Group Phase II Team and the
            FANTOM Consortium.
  TITLE     Functional annotation of a full-length mouse cDNA collection
  JOURNAL   Nature 409, 685-690 (2001)
REFERENCE   5  (bases 1 to 1209)
  AUTHORS   Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A.,
            Arakawa,T., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M.,
            Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F.,
            Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kato,H., Kawai,J.,
            Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T.,
            Miyazaki,A., Nakai,K., [illegible], Nomura,K., [illegible]
            Okido,T., Owa,C., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H.,
            Sasaki,D., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T.,
            Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
            Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A.,
            Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
            Physical and Chemical Research (RIKEN), Laboratory for Genome
            Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
            RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
            Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
            URL:http://[illegible], Tel:81-45-503-[illegible],
            Fax:81-45-503-9216)
COMMENT     Please visit our web site (http://[illegible]) for
            further details.
            cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues. First strand cDNA was primed with a primer
            [5'-GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTVN-3'], cDNA was
            prepared by using trehalose thermo-activated reverse transcriptase
            and subsequently enriched for full-length by cap-trapper. Second
            strand cDNA was prepared with the primer adapter of sequence[5'-
            GAGAGAGAGAGCGGCCGCAATTAATTCTGAGTTAATTAAATTAATCCCCCCCCCCCC-3'], cDNA
            was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3'
            end: SstI. Host: SOLR.
FEATURES             Location/Qualifiers
     source          1..1209
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /db_xref="MGD:MGI:[illegible]"
                     /db_xref="MGD:MGI:[illegible]"
                     /clone="1700012M14"
                     /sex="male"
                     /tissue_type="testis"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="adult"
     CDS             218..1120
                     /note="putative"
                     /codon_start=1
                     /protein_id="BAB24320.1"
                     /db_xref="GI:12838760"
                     /translation="MEE[illegible]YMEVFE[illegible]AASVG
                     DIPQVQKMLEEGD[illegible]DVNVIDKKKRTALHYACAHG[illegible]EMV[illegible]LLWYDCNIEAKDKE[illegible]
                     ALIKAIQ[illegible]HE[illegible]VKILLENGADSNAVD[illegible]HQNIS[illegible]HYIVYNKD[illegible]IAAKL[illegible]AENAD[illegible]
                     VKIKN[illegible]YIPLIIAVLENK[illegible]EMVELLL[illegible]AAAN[illegible]NA[illegible]LN[illegible]KE[illegible]AL[illegible]HAV[illegible]KNH[illegible]SLL
                     LQQGANASLVDIYGATAQSYAVFE[illegible]FQVLSQGPGPKE[illegible]IKEAKEE[illegible]KD[illegible]LV[illegible]
                     PREKEDNKRMEEN[illegible]AVP[illegible]LCRYN"
BASE COUNT      332 a    288 c    328 g    261 t
ORIGIN

  Query Match             17.8%;  Score [illegible];  DB 12;  Length 1209;
  Best Local Similarity   56.5%;  Pred. No. 3.1e-39;
  Matches  381;  Conservative    0;  Mismatches  293;  Indels    0;  Gaps    0;

Qy        391 cccaggtaccacgtccgtggagaagatctggacaagctccacacaac[illegible]ggtgggat 4[illegible]
Db        290 CCCAGGTACCTCACTGGGTACATGCCTGTGAAAAGGATCCACAAAGCTGCAAGCGTGGG[illegible] 34[illegible]

Qy        451 aaagtccccagaaaggatctcatcgtcatgctcaggacactgacgtga[illegible]aagaagac [illegible]
Db        350 GACATACCCCAAGTGCAGAAGATGCTGGAATTGGGGATATTGACGTGAA[illegible]TACAGAC 4[illegible]

Qy        511 aagcaaaagaggactgctctacatctggcctctgccaatgggaat[illegible]cagaagtagtaaaa [illegible]
Db        410 CGGAAGAAAAGGACTGCTCTTCACTACGCCTGTGCTCACGGCCAGTCAGAAATGGTGAGC 4[illegible]

Qy        571 ctcctgctggacagacgatgtcaacttaatgtccttgacaacaaaaaagacagctctg [illegible]
Db        47[illegible] [illegible]

Qy        631 ataaaggccgtacaatgccaggaagatgaatgtgcgttaatgttgctggaacatggcact [illegible]
Db        530 ATCAAGGC[illegible]ACACAGC[illegible]AG[illegible]A[illegible]GAGATCTGTGTGAAGATTCTT[illegible]A[illegible]AAC[illegible] 5[illegible]

Qy        691 gatccaaatattccagatgagtatggaaataccactctgcactacgc[illegible]atctataatgaa 7[illegible]0
Db        590 GACTCCAACGCCGTCGATATCCATCAGAACAC[illegible]CTCCACTACAC[illegible]TACAACAAC [illegible]

Qy        751 gataaattaatggccaaagcactgctcttatatggtgctgatatcaaatcaaaaacaag 8[illegible]
Db        [illegible]

Qy        8[illegible]1 [illegible]actgttacttggtgtacatgagcaaaacag[illegible]aagt[illegible]gtgaaattt 8[illegible]
Db        710 AACGGCTACACG[illegible]GCTTATACTTGCGG[illegible]TAGAAACAAACA[illegible]A[illegible]AT[illegible]GAAC[illegible] 7[illegible]

Qy        871 ttaatcaagaaaaaagcgaatttaaatgcactggataggtatggaaggactgctctcat[illegible] 9[illegible]
Db        770 TTACTGCAGGCTGCAGCCAACATAAACGCCCTGGATAACTGTAACAG[illegible]TTATA 8[illegible]

Qy        931 cttgctgtatgttgtggatcagcaagtatagtcagccttctact[illegible]aaatattgat [illegible]
Db        830 CATGCTGTGAGAACTCAGTCTAAAAACATCATCAGC[illegible]TCTTCTCCA[illegible]AAGGC[illegible]AAT 8[illegible]

Qy        991 gtatcttctcaagatctatctggacagacgaccagagagtatgct[illegible]tagtcatgat [illegible]
Db        890 GCATCTTTGGTGGACATCTATGGAGCTACTGCACAAAGTTACGCTGT[illegible]TTGAAACTTT 94[illegible]

Qy       1051 catgtaatttgcca 1064
Db        950 CAAGTGCTTTCCCA 963
```

Search completed: March 9, 2002, 01:50:43
Job time: 4922 sec

```
Nucleic - nucleic search, using sw model

Run on:          March 9, 2002, 03:18:10 ; Search time 156.99 Seconds
                                            (without alignments)
                                            6307.476 Million cell updates/sec

Title:           US-09-699-295-301
Perfect score:   1155
Sequence:        1 atqqtqqttqaqqttqattc.........caaqaaccaqaaataaataa 1155

Scoring table:   OLIGO NUC
                 Gapop 60.0 , Gapext 60.0

Searched:        930621 seqs, 428662619 residues

Word size :      9

Total number of hits satisfying chosen parameters:      1861242

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :       N_Geneseq_1101:*
                 1: [illegible]NA1980.DAT:*
                 2: [illegible]NA1981.DAT:*
                 3: [illegible]NA1982.DAT:*
                 4: [illegible]NA1983.DAT:*
                 5: [illegible]NA1984.DAT:*
                 6: [illegible]NA1985.DAT:*
                 7: [illegible]NA1986.DAT:*
                 8: [illegible]NA1987.DAT:*
                 9: [illegible]NA1988.DAT:*
                 10: [illegible]NA1989.DAT:*
                 11: [illegible]NA1990.DAT:*
                 12: [illegible]NA1991.DAT:*
                 13: [illegible]NA1992.DAT:*
                 14: [illegible]NA1993.DAT:*
                 15: [illegible]NA1994.DAT:*
                 16: [illegible]NA1995.DAT:*
                 17: [illegible]NA1996.DAT:*
                 18: [illegible]NA1997.DAT:*
                 19: [illegible]NA1998.DAT:*
                 20: [illegible]NA1999.DAT:*
                 21: [illegible]NA[illegible].DAT:*
                 22: [illegible]NA[illegible].DAT:*

   Pred. No. is the number of results predicted by chance to have a
   score greater than or equal to the score of the result being printed,
   and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		1155	100.0	1155	21	AAC81011	Human B11Aq1 antiq
2		1155	100.0	1155	21	AAA06598	Human immunogenic
3		1155	100.0	1155	22	AAH93714	Human prostate spe
4		1155	100.0	1155	22	AAH85028	Human prostate spe
5		1155	100.0	1155	22	AAH02779	Prostate tumour an
6		1128	97.7	2000	21	AAC81012	Human B11Aq1 antiq
7		1128	97.7	2000	22	AAH93715	Human prostate spe
8		1128	97.7	2000	22	AAH85029	Human prostate spe
9		1128	97.7	2000	22	AAH02780	Prostate tumour an
10		1128	97.7	2040	21	AAC81013	Human B11Aq1 antiq
11		1128	97.7	2040	21	AAA06600	Human immunogenic
12		1128	97.7	2040	22	AAH93716	Human prostate spe
13		1128	97.7	2040	22	AAH85030	Human prostate spe
14		1128	97.7	2040	22	AAH02781	Prostate tumour an
15		1077	93.2	2000	21	AAA06599	Human immunogenic
16		491	42.5	1512	19	AAV68995	DNA molecule encod
17		491	42.5	1512	21	AAC81006	Human B11Aq1 antiq
18		491	42.5	1512	21	AAA06593	Human immunogenic
19		491	42.5	1512	22	AAH93709	Human prostate spe
20		491	42.5	1512	22	AAH85023	Human prostate spe
21		491	42.5	1512	22	AAH02774	Prostate tumour an
22		252	21.8	1853	19	AAV68996	DNA molecule encod
23		252	21.8	1853	21	AAC81007	Human B11Aq1 antiq
24		252	21.8	1853	21	AAA06594	Human immunogenic
25		252	21.8	1853	22	AAH93710	Human prostate spe
26		252	21.8	1853	22	AAH85024	Human prostate spe
27		252	21.8	1853	22	AAH02775	Prostate tumour an
28		120	10.4	879	21	AAC81016	Human breast tumou
29		120	10.4	879	22	AAH93826	Human prostate spe
30		120	10.4	879	22	AAH85140	Human prostate spe
31		120	10.4	1059	21	AAC81010	Human B11Aq1 antiq
32		120	10.4	1059	21	AAA06597	Human immunogenic
33		120	10.4	1059	22	AAH93713	Human prostate spe
34		120	10.4	1059	22	AAH85027	Human prostate spe
35		120	10.4	1059	22	AAH02778	Prostate tumour an
36	c	120	10.4	1851	19	AAV68992	DNA molecule encod
37	c	120	10.4	1851	19	AAV68993	DNA molecule encod
38	c	120	10.4	1851	21	AAC81003	Human breast tumou
39	c	120	10.4	1851	21	AAC81004	Human B11Aq1 antiq
40	c	120	10.4	1851	21	AAA06591	Human immunogenic
41	c	120	10.4	1851	22	AAH93707	Human prostate spe
42	c	120	10.4	1851	22	AAH85021	Human prostate spe
43	c	120	10.4	1851	22	AAH02772	Prostate tumour an
44		120	10.4	1852	21	AAC81015	Human breast tumou
45		120	10.4	1852	22	AAH93825	Human prostate spe

ALIGNMENTS

```
RESULT 1
AAC81011
ID   AAC81011 standard; cDNA; 1155 BP.
XX
AC   AAC81011;
XX
DT   15-FEB-2001  (first entry)
XX
DE   Human B11Aq1 antigen splice variant B11C-15 cDNA.
XX
KW   Human; breast tumour; specific antigen; cytostatic; vaccine;
KW   breast cancer; B18Aq1; B11Aq1; B15Aq1; ss.
XX
OS   Homo sapiens.
XX
PN   WO200061753-A2.
XX
PD   19-OCT-2000.
XX
PF   07-APR-2000; 2000WO-US09312.
XX
PR   09-APR-1999;   99US-0289198.
PR   28-OCT-1999;   99US-0429755.
PR   23-MAR-2000; 2000US-0534825.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Frudakis TN,  Smith JM,  Reed SG,  Misher LE,  Retter MW,  Dillon DC
XX
DR   WPI; 2000-628403/60.
DR   P-PSDB; AAB28628.
XX
PT   An isolated polypeptide comprising an immunogenic portion of a breast
PT   tumor protein used for inhibiting the development of cancer, especially
```

PT breast cancer, and monitoring cancer progression in a patient -
XX
PS Claim 4; Page 177; 187pp; English.
XX
CC The present sequence is given in a specification relating to compositions
CC and methods for the treatment and diagnosis of breast cancer. Nucleotide
CC sequences that are preferentially expressed in breast tumour tissue, and
CC the polypeptides encoded by such nucleotide sequences, are used in
CC compositions and vaccines to inhibit the development of cancer,
CC especially breast cancer. The progression of a cancer may be monitored by
CC carrying out detection of tumour specific antigens at subsequent time
CC points and comparing the results from the different time points.
CC CD4+ and/or CD8+ T-Cells isolated from the cancer patient may be treated
CC with tumour-specific polypeptides, polynucleotides encoding the
CC polypeptides or antigen presenting cells expressing the polypeptides. The
CC cells are then administered to the patient to inhibit development of
CC cancer.
XX
SQ Sequence 1155 BP; 346 A; 253 C; 297 G; 259 T; 0 other;

Query Match 100.0%; Score 1155; DB 21; Length 1155;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1155; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 atggtgqttgaqqttgattccatqccgqctgcctcttctqtgaagaagccatttggtctc 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 atggtgqttgaqqttgattccatqccgqctgcctcttctqtgaagaagccatttggtctc 60

Qy      61 aqgagcaagatgqgcaagtggtgctgccgttgcttccccctgctgcagggagagcggcaag 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 aggagcaagatgggcaagtggtgctgccgttgcttccccctgctgcagggagagcggcaag 120

Qy     121 agcaagatgagcagttctgggagaccagcagacgactctgctatgaagacactcaggagcaag 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 agcaagtgggcagttctgggagaccagcagacgactctgctatgaagacactcaggagcaag 180

Qy     181 atgggcaagtggtgccgccactgcttccctgctgcaggggagtggcaagagcaagcgtg 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 atgggcaagtggtgccgccactgcttccctgctgcaggggagtggcaagagcaagcgtg 240

Qy     241 qgcgcttctgqaqaccagcagcgactctgctatgaagacactcaggaacaagatgtggcaag 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 qgcgcttctggagaccagcagcgactctgctatgaagacactcaggaacaagatgtggcaag 300

Qy     301 tggtgctgccactgcttccctgctgcagggggaqcggcaagagcaaggtgggcgcttgg 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 tggtgctgccactgcttccctgctgcagggggaqcggcaagagcaaggtgggcgcttgg 360

Qy     361 qaagactacgatgacaagtgccttatgaagccaaggtaccacgtccgtggagaaqatctg 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 qaagactacgatgacaagtgccttcatgaagccaaggtaccacgtccgtggagaaqatctg 420

Qy     421 qacaaqctccacaqaqctgcctgtgtgqqqtaaaqtcccqagaaaggatctcatcgtcatg 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 qacaaqctccacaqaqctgcctgtgtgqqqtaaaqtcccqagaaaggatctcatcgtcatg 480

Qy     481 ctcaqqqacactqacgtqaacaaqqaaqqacaaqcaaaaqaggactgctctacatctggcc 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 ctcaqqqacactqacgtqaacaaqqaaqqacaaqcaaaaqaggactgctctacatctggcc 540

Qy     541 tctgccaatgqqaattcagaaqtagtaaaactcctgctggacagacgatgtcaacttaat 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 tctgccaatgqqaattcagaaqtagtaaaactcctgctggacagacgatgtcaacttaat 600

Qy     601 qtccttqacaacaaaaaqaggacagctctgataaaggccgtacaatgccaggaagatgaa 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 qtccttqacaacaaaaaqaggacagctctgataaaggccgtacaatgccaggaagatgaa 660

Qy     661 tgtgcgttaatgttgctgtggaacatggcactgatccaaatattccagatgagtatggaaat 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 tgtgcgttaatgttgctgtggaacatggcactgatccaaatattccagatgagtatggaaat 720

Qy     721 accactctgcactacgctatctataatgaagataaattaatggccaaagcactactcttta 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 accactctgcactacgctatctataatgaagataaattaatggccaaagcactactcttta 780

Qy     781 tatggtgctgatatcgaatcaaaaaacaagcatgcctctcaccactgttactttqtatta 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 tatggtgctgatatcgaatcaaaaaacaagcatgcctctcaccactgttactttqtatta 840

Qy     841 catgaqcaaaaacagcaaqtcgtgaaattttaatcaaqaaaaaaqcgaattaaatgca 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 catgagcaaaaacagcaagtcgtgaaattttaatcaagaaaaaaqcgaattaaatgca 900

Qy     901 ctggataqatatqgaaggactgctctcatacttgctgtatgttgtggatcaqcaagtata 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 ctggataqatatqgaaggactgctctcatacttgctgtatgttgtggatcagcaagtata 960

Qy     961 gtcagccttctacttgagcaaaatattgatgtatcttctcaagatctatctggacagacg 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 gtcagccttctacttgagcaaaatattgatgtatcttctcaagatctatctggacagacg 1020

Qy    1021 gccagagaqtatgctgtttctaqtcatcatcatgtaatttgccaqttacttttctgactac 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 gccagagagtatgctgtttctagtcatcatcatgtaatttgccaqttacttttctgactac 1080

Qy    1081 aaaqaaaaacagatgctaaaaatctcttctqaaaaacagcaatccaqaaaatgtctcaaga 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 aaaqaaaaacagatgctaaaaatctcttctqaaaaacagcaatccaqaaaatgtctcaaga 1140

Qy    1141 accagaaataaataa 1155
          ||||||||||||||||
Db    1141 accagaaataaataa 1155
```

RESULT 3
AAA06598
ID AAA06598 standard; cDNA; 1155 BP.
XX
AC AAA06598;
XX
DT 14-JUN-2000 (first entry)
XX
DE Human immunogenic prostate tumour protein cDNA sequence SEQ ID NO:373.
XX
KW Human; prostate cancer; diagnosis; tumour; gene therapy; detection;
KW immunogenic; cytostatic; vaccine; ss.
XX
OS Homo sapiens.
XX
PN WO200004149-A2.
XX
PD 27-JAN-2000.
XX
PF 14-JUL-1999; 99WO-US15838.
XX
PR 14-JUL-1998; 98US-0115453.
PR 14-JUL-1998; 98US-0116134.
PR 23-SEP-1998; 98US-0159812.
PR 23-SEP-1998; 98US-0159822.
PR 15-JAN-1999; 99US-0232149.
PR 15-JAN-1999; 99US-0232880.
PR 09-APR-1999; 99US-0288946.
XX
PA (CORI-) CORIXA CORP.
XX
PI Dillon DC, Harlocker SL, Yuqiu J, Xu J, Mitcham JL;
XX
DR WPI; 2000-171268/15.
XX
PT New polypeptide useful for treating and diagnosing prostate cancer
PT comprises an immunogenic portion of prostate tumor protein -
XX
PS Claim 50; Page 222; 263pp; English.
XX

```
CC   The present invention describes isolated polypeptides, comprising an
CC   immunogenic portion of a prostate tumour protein (PTP). The polypeptides
CC   and polynucleotides encoding them have cytostatic activity and can be
CC   used in vaccines and in gene therapy. The polypeptides and
CC   polynucleotides encoding them, antigen presenting cells which express
CC   the polypeptides, antibodies against the polypeptides and vaccines
CC   comprising them can be used for inhibiting the development of prostate
CC   cancer in a patient. The polypeptides can be used to generate antibodies
CC   or anti-idiotypic antibodies for passive immuno therapy. A portion of
CC   the polynucleotides encoding the polypeptides can be used as a probe or
CC   to modulate the expression of the polypeptides. AAA06241 to AAA06691 and
CC   AAY82800 to AAY82020 represent sequences used in the exemplification of
CC   the present invention.
XX
SQ   Sequence 1155 BP; 346 A; 253 C; 297 G; 259 T; 0 other;

  Query Match             100.0%;  Score 1155;  DB 21;  Length 1155;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1155;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 [illegible]gattccatgccgggtgc[illegible]cttctgtgaagaagccatttggtctc 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 [illegible]gattccatgccgggtgc[illegible]cttctgtgaagaagccatttggtctc 60

Qy         61 [illegible]ggcaagtggtgctgccgtttg[illegible]ttcccctgctgcaggggaagcggcaag 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 [illegible]ggcaagtggtgctgccgtttg[illegible]ttcccctgctgcaggggaagcggcaag 120

Qy        121 [illegible]tggggacttctggagaccacgacgactctgctatgaagacactcaggagcaag 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 [illegible]tggggacttctggagaccacgacgactctgctatgaagacactcaggagcaag 180

Qy        181 atgggcaagtggtgccgccactgcttcccctgctgcaggggaagtggcaagagcaacgtg 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 atgggcaagtggtgccgccactgcttcccctgctgcaggggaagtggcaagagcaacgtg 240

Qy        241 gccgcttctggagaccacgacgactctgctatgaagacactcaggaacaagatgggcaag 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 gccgcttctggagaccacgacgactctgctatgaagacactcaggaacaagatgggcaag 300

Qy        301 t[illegible]ctgccactgcttcccctgctgcagggggagcggcaagagcaaggtgggcgcttgg 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 t[illegible]ctgccactgcttcccctgctgcagggggagcggcaagagcaaggtgggcgcttgg 360

Qy        361 [illegible]ctacgatgacagtgccttcatggagcccaggtaccacgtccgtggagaagatctg 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 [illegible]ctacgatgacagtgccttcatggagcccaggtaccacgtccgtggagaagatctg 420

Qy        421 [illegible]ctccacagagctgcctggtgggg[illegible]agtccccaggaaggatctcatcgtcatg 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 [illegible]ctccacagagctgcctggtgggg[illegible]agtccccaggaaggatctcatcgtcatg 480

Qy        481 [illegible]tgacagtgaacaaaggacaagcaaaaggaggactgctctacatctggcc 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 [illegible]tgacagtgaacaaaggacaagcaaaaggaggactgctctacatctggcc 540

Qy        541 tctgccaatggggaattcagaagtagtaaaactcctgctggacagacgatgtcaacttaat 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 tctgccaatggggaattcagaagtagtaaaactcctgctggacagacgatgtcaacttaat 600

Qy        601 gtccttgacaacaaaaagaggacagctctgataaaggccgtacaatgccaggaagatgaa 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 gtccttgacaacaaaaagaggacagctctgataaaggccgtacaatgccaggaagatgaa 660

Qy        661 tgtgcgttaatgttgctggaacatggcactgatccaaatattccagatgagtatggaaat 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 tgtgcgttaatgttgctggaacatggcactgatccaaatattccagatgagtatggaaat 720

Qy        721 [illegible]ctatctataatgaagataattaatggccaaagcactgctctta 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 [illegible]ctatctataatgaagataattaatggccaaagcactgctctta 780

Qy        781 tatggtgctgatatcgaatcaaaaaacaagcatggcctcacacccactgttacttggtgta 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 tatggtgctgatatcgaatcaaaaaacaagcatggcctcacacccactgttacttggtgta 840

Qy        841 catgagcaaaaacagcaagtcgtgaaatttttaatcaagaaaaaa[illegible]gaatttaaatgca 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 catgagcaaaaacagcaagtcgtgaaatttttaatcaagaaaaaa[illegible]gaatttaaatgca 900

Qy        901 ctggatagatatggaaggactgctctcatacttgctgtatgttgtgga[illegible]aagtata 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 ctggatagatatggaaggactgctctcatacttgctgtatgttgtgga[illegible]aagtata 960

Qy        961 gtcagccttctacttgagcaaaatattgatgtatcttctcaagatctatctggacagacg 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 gtcagccttctacttgagcaaaatattgatgtatcttctcaagatctatctggacagacg 1020

Qy       1021 gccagagagtatgctgtttctagtcatcatcatgtaatttgccagttatttttctgactac 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 gccagagagtatgctgtttctagtcatcatcatgtaatttgccagttatttttctgactac 1080

Qy       1081 aaagaaaaacagatgctaaaatctcttctgaaaacagcaatcc[illegible]tctcaaga 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 aaagaaaaacagatgctaaaatctcttctgaaaacagcaatcc[illegible]tctcaaga 1140

Qy       1141 accagaaataaataa 1155
              |||||||||||||||
Db       1141 accagaaataaataa 1155


RESULT 3
AAH93714
ID   AAH93714 standard; cDNA; 1155 BP.
XX
AC   AAH93714;
XX
DT   04-OCT-2001  (first entry)
XX
DE   Human prostate specific cDNA sequence B305D splice variant #8.
XX
KW   Human; prostate cancer; prostate specific; diagnosis; vaccine;
KW   cytostatic; gene therapy; metastasis; ss.
XX
OS   Homo sapiens.
XX
PN   WO200151633-A2.
XX
PD   19-JUL-2001.
XX
PF   16-JAN-2001; 2001WO-US01574.
XX
PR   14-JAN-2000; 2000US-0483672.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Xu J,  Dillon DC,  Mitcham JL,  Harlocker SL,  Jiang Y,  Reed SG,
PI   Kalos MD,  Fanger GR,  Day CH,  Retter MW,  Stolk JA,  Skeiky YAW;
PI   Wang A,  Meagher MJ;
XX
DR   WPI; 2001-425873/45.
XX
PT   New polynucleotide encoding a prostate specific protein, for
PT   diagnosing, monitoring and treating prostate cancer in a patient and
PT   for use in vaccines
XX
PS   Claim 1; Page 347; 543pp; English.
XX
CC   The present invention describes polynucleotide sequences (I) which encode
CC   prostate specific proteins (II). (I) and (II) have cytostatic activity
CC   and can be used in vaccine production and gene therapy. (I), (II),
CC   antibodies to (II), fusion proteins comprising (II), and isolated
CC   T cells prepared using (I) or (II) are used treat cancer in a patient
CC   (I) and the antibodies are also used in the detection of cancer in a
CC   patient. The cancer that is diagnosed or treated is particularly
```

```
CC   prostate cancer  (I) and (II) can be used in vaccines. The antibodies or
CC   (I) can be used for monitoring the progression of cancer in a patient.
CC   (I) and (II) can also be used to improve diagnostic and therapeutic
CC   methods for prostate cancer. They can indicate the level of metastasis
CC   as well as the prostate volume. AAH93357 to AAH93944 and AAM01115 to
CC   AAM01318 represent polynucleotide and amino acid sequences used in the
CC   exemplification of the present invention.
XX
SQ   Sequence 1155 BP; 346 A; 253 C; 297 G; 259 T; 0 other

  Query Match             100.0%;  Score 1155;  DB 22;  Length 1155;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1155;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 atggtggttgaggttgattccatgccggctgcctcttctgtgaagaagccatttggtctc 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 atggtggttgaggttgattccatgccggctgcctcttctgtgaagaagccatttggtctc 60

Qy         61 aggagcaagatgggcaagtggtgctgccgttgcttcccctgctgcagggagagcggcaag 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 aggagcaagatgggcaagtggtgctgccgttgcttcccctgctgcagggagagcggcaag 120

Qy        121 agcaacgtgggcacttctggagacccacgacgactctgctatgaagacactcaggagcaag 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 agcaacgtgggcacttctggagacccacgacgactctgctatgaagacactcaggagcaag 180

Qy        181 atgggcaagtggtgccgccactgcttcccctgctgcagggggagtggcaagagcaacgtg 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 atgggcaagtggtgccgccactgcttcccctgctgcagggggagtggcaagagcaacgtg 240

Qy        241 ggcgcttctggagacccacgacgactctgctatgaagacactcaggaacaagatgggcaag 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ggcgcttctggagacccacgacgactctgctatgaagacactcaggaacaagatgggcaag 300

Qy        301 tggtgctgccactgcttcccctgctgcagggggagcggcaagagcaaggtgggcgcttgg 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 tggtgctgccactgcttcccctgctgcagggggagcggcaagagcaaggtgggcgcttgg 360

Qy        361 ggagactacgatgacagtgccttcatggagcccaggtaccacgtccgtggagaagatctg 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ggagactacgatgacagtgccttcatggagcccaggtaccacgtccgtggagaagatctg 420

Qy        421 gacaagctccacagagctgcctggtggggtaaagtccccagaaaggatctcatcgtcatg 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 gacaagctccacagagctgcctggtggggtaaagtccccagaaaggatctcatcgtcatg 480

Qy        481 ctcagggacactgacgtgaacaagaaggacaagcaaaagaggactgctctacatctggcc 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ctcagggacactgacgtgaacaagaaggacaagcaaaagaggactgctctacatctggcc 540

Qy        541 tctgccaatgggaattcagaagtagtaaaactcctgctggacagacgatgtcaacttaat 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 tctgccaatgggaattcagaagtagtaaaactcctgctggacagacgatgtcaacttaat 600

Qy        601 gtccttgacaacaaaaagaggacagctctgataaaggccgtacaatgccaggaagatgaa 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 gtccttgacaacaaaaagaggacagctctgataaaggccgtacaatgccaggaagatgaa 660

Qy        661 tgtgcgttaatgttgctggaacatggcactgatccaaatattccagatgagtatggaaat 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 tgtgcgttaatgttgctggaacatggcactgatccaaatattccagatgagtatggaaat 720

Qy        721 accactctgcactacgctatctataatgaagataaattaatggccaaagcactgctctta 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 accactctgcactacgctatctataatgaagataaattaatggccaaagcactgctctta 780

Qy        781 tatggtgctgatatcgaatcaaaaaacaagcatggcctcacaccactgttacttggtgta 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 tatggtgctgatatcgaatcaaaaaacaagcatggcctcacaccactgttacttggtgta 840

Qy        841 catgagcaaaaacagcaagtcgtgaaatttttaatcaagaaaaaagcgaatttaaatgca 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 catgagcaaaaacagcaagtcgtgaaatttttaatcaagaaaaaagcgaatttaaatgca 900

Qy        901 ctggatagatatggaaggactgctctcatacttgctgtatgttgtggatcagcaagtata 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 ctggatagatatggaaggactgctctcatacttgctgtatgttgtggatcagcaagtata 960

Qy        961 gtcagccttctacttgagcaaaatattgatgtatcttctcaagatctatctggacagacg 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 gtcagccttctacttgagcaaaatattgatgtatcttctcaagatctatctggacagacg 1020

Qy       1021 gccagagagtatgctgtttctagtcatcatcatgtaatttgccagttactttctgactac 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 gccagagagtatgctgtttctagtcatcatcatgtaatttgccagttactttctgactac 1080

Qy       1081 aaagaaaaacagatgctaaaaatctcttctgaaaacagcaatccagaaaatgtctcaaga 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 aaagaaaaacagatgctaaaaatctcttctgaaaacagcaatccagaaaatgtctcaaga 1140

Qy       1141 accagaaataaataa 1155
              |||||||||||||||
Db       1141 accagaaataaataa 1155


RESULT  4
AAH85028
ID   AAH85028 standard; cDNA; 1155 BP.
XX
AC   AAH85028;
XX
DT   25-SEP-2001  (first entry)
XX
DE   Human prostate specific cDNA sequence B305D splice variant #8.
XX
KW   human; prostate cancer; therapy; diagnosis; cat eye syndrome;
KW   chromosome 22q11.2; prostate specific protein; chromosome 1;
KW   prostate specific antigen; PSA; ss.
XX
OS   Homo sapiens.
XX
PN   WO200134802-A2.
XX
PD   17-MAY-2001.
XX
PF   09-NOV-2000; 2000WO-US30904.
XX
PR   12-NOV-1999;   99US-0439313.
PR   18-NOV-1999;   99US-0443686.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Xu J,  Dillon DC,  Mitcham JL,  Harlocker SL,  Jiang Y,  Reed SG;
PI   Kalos MD,  Retter MW,  Stolk JA,  Day CH,  Skeiky YAW,  Wang A;
XX
DR   WPI; 2001-398785/32.
XX
PT   Isolated polypeptide comprising at least an immunogenic portion of a
PT   prostate-specific protein, useful in the diagnosis and therapy of
PT   prostate cancer -
XX
PS   Claim 31; Page 246-247; 325pp; English.
XX
CC   The present invention describes an isolated polypeptide (I) comprising
CC   at least an immunogenic portion of a prostate-specific protein, or its
CC   variant. Also described are polynucleotides (N1) encoding (I). (P1) and
CC   (N1) have cytostatic activity and can be used in vaccine production.
CC   The polypeptides, nucleic acids and antibodies from the present
CC   invention are useful in the diagnosis and therapy of prostate cancer.
CC   Prostate specific genes P704P, P712P, P774P, P775P and P501S are located
CC   in a genomic region on chromosome 22q11.2 known as the Cat Eye Syndrome
CC   region. Prostate specific antigen (PSA) P501S was located on
CC   chromosome 1. AAH84671 to AAH85143 and AAG99000 to AAG99077 represent
CC   polynucleotide and polypeptide sequences used in the exemplification
CC   of the present invention.
```

```
XX
SQ   Sequence 1155 BP; 346 A; 253 C; 297 G; 259 T; 0 other;

  Query Match             100.0%;  Score 1155;  DB 22;  Length 1155;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1155;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 atgatggttgaggttgattccatgccggctgcctcttctgtgaagaagccatttggtctc 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 atgatggttgaggttgattccatgccggctgcctcttctgtgaagaagccatttggtctc 60

Qy         61 attaggcaaatggacaagtggtgctgccgttgcttcccctgctgcaggggagaggcggcaag 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 attaggcaaatggacaagtggtgctgccgttgcttcccctgctgcaggggagaggcggcaag 120

Qy        121 agcaggttggcacttctggagaccacgacgactctgctatgaagacactcaggagcaag 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 agcaggttggcacttctggagaccacgacgactctgctatgaagacactcaggagcaag 180

Qy        181 atgggcaagtggtgccgccactgcttcccctgctgcaggggagtggcaagagcaacgtg 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 atgggcaagtggtgccgccactgcttcccctgctgcaggggagtggcaagagcaacgtg 240

Qy        241 agcaggcttctggagaccacgacgactctgctatgaagacactcaggaacaagatgggcaag 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 agcaggcttctggagaccacgacgactctgctatgaagacactcaggaacaagatgggcaag 300

Qy        301 tggtgctgccactgcttcccctgctgcagggggagcggcaagagcaaggtgggcgcttgg 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 tggtgctgccactgcttcccctgctgcagggggagcggcaagagcaaggtgggcgcttgg 360

Qy        361 gggaactacgatgacagtgccttcatggagcccaggtaccacgtccgtggagaagatctg 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 gggaactacgatgacagtgccttcatggagcccaggtaccacgtccgtggagaagatctg 420

Qy        421 gacaagctccacagagctgcctggtggggtaaagtcccagaaaggatctcatcgtcatg 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 gacaagctccacagagctgcctggtggggtaaagtcccagaaaggatctcatcgtcatg 480

Qy        481 ctcagggacactgacgtgaacaagaaggacaagcaaaagaggactgctctacatctggcc 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ctcagggacactgacgtgaacaagaaggacaagcaaaagaggactgctctacatctggcc 540

Qy        541 tctgccaatgggaattcagaagtagtaaaactcctgctggacagacgatgtcaacttaat 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 tctgccaatgggaattcagaagtagtaaaactcctgctggacagacgatgtcaacttaat 600

Qy        601 gtccttgacaacaaaaagaggacagctctgataaaggccgtacaatgccaggaagatgaa 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 gtccttgacaacaaaaagaggacagctctgataaaggccgtacaatgccaggaagatgaa 660

Qy        661 tgtgcgttaatgttgctggaacatggcactgatccaaatattccagatgagtatggaaat 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 tgtgcgttaatgttgctggaacatggcactgatccaaatattccagatgagtatggaaat 720

Qy        721 accactctgcactacgctatctataatgaagataaattaatggccaaagcactgctctta 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 accactctgcactacgctatctataatgaagataaattaatggccaaagcactgctctta 780

Qy        781 tatggtgctgatatcgaatcaaaaaacaagcatggcctcacaccactgttacttggtgta 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 tatggtgctgatatcgaatcaaaaaacaagcatggcctcacaccactgttacttggtgta 840

Qy        841 catgagcaaaacagcaagtcgtgaaatttttaatcaagaaaaaagcgaatttaaatgca 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 catgagcaaaacagcaagtcgtgaaatttttaatcaagaaaaaagcgaatttaaatgca 900

Qy        901 ctggatagatatggaaggactgctctcatacttgctgtatgttgtggatcagcaagtata 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 ctggatagatatggaaggactgctctcatacttgctgtatgttgtggatcagcaagtata 960

Qy        961 gtcagccttctacttgagcaaaatattgatgtatcttctcaagatctatctggacagacg 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 gtcagccttctacttgagcaaaatattgatgtatcttctcaagatctatctggacagacg 1020

Qy       1021 gccagagagtatgctgtttctagtcatcatcatgtaatttgccagttactttctgactac 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 gccagagagtatgctgtttctagtcatcatcatgtaatttgccagttactttctgactac 1080

Qy       1081 aagagaaaaacagatgctaaaaatctcttctgaaaacagcaatccagaaaatatctcaaa 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 aagagaaaaacagatgctaaaaatctcttctgaaaacagcaatccagaaaatatctcaaa 1140

Qy       1141 accagaaataaataa 1155
              |||||||||||||||
Db       1141 accagaaataaataa 1155


RESULT 5
AAH02779
ID   AAH02779 standard; cDNA; 1155 BP.
XX
AC   AAH02779;
XX
DT   14-JUN-2001  (first entry)
XX
DE   Prostate tumour antigen determined cDNA splice variant of B305D #4
XX
KW   Human; prostate tumour antigen; prostate tumour; therapy; diagnosis;
KW   prostate cancer; immunogenic; cytostatic; vaccine; ss.
XX
OS   Homo sapiens.
XX
PN   WO200125272-A2.
XX
PD   12-APR-2001.
XX
PF   04-OCT-2000; 2000WO-US27464.
XX
PR   04-OCT-1999;   99US-0157455.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Xu J,  Skeiky YAW,  Reed SG,  Cheever MA;
XX
DR   WPI; 2001-245862/25.
DR   P-PSDB; AAB74815.
XX
PT   Prostate specific protein and its encoding polynucleotide, useful for
PT   the treatment and diagnosis of prostate cancer
XX
PS   Claim 50; Page 231-232; 276pp; English.
XX
CC   The present invention describes an isolated polypeptide (I) comprising
CC   at least an immunogenic portion of a prostate tumour antigen protein or
CC   its variant. (I) have cytostatic activity and can be used in vaccine
CC   production. (I), prostate tumour antigen polynucleotides, an antigen
CC   presenting cell (APC e.g. a dendritic cell) that expresses (I), and a
CC   pharmaceutical composition containing (I) are useful for inhibiting the
CC   development of cancer in a patient. Antibodies specific for prostate
CC   specific proteins and oligonucleotides that hybridise to a
CC   polynucleotide that encodes a prostate specific protein are useful
CC   for detecting the presence or absence of a cancer or monitoring the
CC   progression the progression of a cancer, especially prostate cancer.
CC   AAH02422 to AAH2872, AAB74708 to AAB74821 and AAB74906 are sequences
CC   used in the exemplification of the present invention.
XX
SQ   Sequence 1155 BP; 346 A; 253 C; 297 G; 259 T; 0 other;

  Query Match             100.0%;  Score 1155;  DB 22;  Length 1155;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1155;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 atgatggttgaggttgattccatgccggctgcctcttctgtgaagaagccatttggtctc 60
```

```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 atggtggttgaggttgattccatgccggctgcctcttctgtgaagaagccatttggtctc 60

Qy         61 aggagcaagatgggcaagtggtgctgccgttgcttcccctgctgcagggagagcggcaag 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 aggagcaagatgggcaagtggtgctgccgttgcttcccctgctgcagggagagcggcaag 120

Qy        121 agcaacgtgggcacttctggagaccacgacgactctgctatgaagacactcaggagcaag 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 agcaacgtgggcacttctggagaccacgacgactctgctatgaagacactcaggagcaag 180

Qy        181 atgggcaagtggtgccgccactgcttcccctgctgcagggggagtggcaagagcaacgtg 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 atgggcaagtggtgccgccactgcttcccctgctgcagggggagtggcaagagcaacgtg 240

Qy        241 ggcgcttctggagaccacgacgactctgctatgaagacactcaggaacaagatgggcaag 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ggcgcttctggagaccacgacgactctgctatgaagacactcaggaacaagatgggcaag 300

Qy        301 tggtgctgccactgcttcccctgctgcagggggagcggcaagagcaaggtgggcgcttgg 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 tggtgctgccactgcttcccctgctgcagggggagcggcaagagcaaggtgggcgcttgg 360

Qy        361 ggagactacgatgacagtgccttcatggaggccaggtaccacgtccgtggagaagatctg 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ggagactacgatgacagtgccttcatggaggccaggtaccacgtccgtggagaagatctg 420

Qy        421 gacaagctccacagagctgcctggtggggtaaagtccccagaaggatctcatcgtcatg 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 gacaagctccacagagctgcctggtggggtaaagtccccagaaggatctcatcgtcatg 480

Qy        481 ctcagggacactgacgtgaacaagaaggacaagcaaaagaggactgctctacatctggcc 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ctcagggacactgacgtgaacaagaaggacaagcaaaagaggactgctctacatctggcc 540

Qy        541 tctgccaatgggaattcagaagtagtaaaactcctgctggacagacgatgtcaacttaat 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 tctgccaatgggaattcagaagtagtaaaactcctgctggacagacgatgtcaacttaat 600

Qy        601 gtccttgacaacaaaaagaggacagctctgataaaggccgtacaatgccaggaagatgaa 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 gtccttgacaacaaaaagaggacagctctgataaaggccgtacaatgccaggaagatgaa 660

Qy        661 tgtgcgttaatgttgctggaacatggcactgatccaaatattccagatgagtatggaaat 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 tgtgcgttaatgttgctggaacatggcactgatccaaatattccagatgagtatggaaat 720

Qy        721 accactctgcactacgctatctataatgaagataaattaatggccaaagcactgctctta 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 accactctgcactacgctatctataatgaagataaattaatggccaaagcactgctctta 780

Qy        781 tatggtgctgatatcgaatcaaaaaacaagcatggcctcacaccactgttacttggtgta 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 tatggtgctgatatcgaatcaaaaaacaagcatggcctcacaccactgttacttggtgta 840

Qy        841 catgagcaaaaacagcaagtcgtgaaatttttaatcaagaaaaaagcgaatttaaatgca 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 catgagcaaaaacagcaagtcgtgaaatttttaatcaagaaaaaagcgaatttaaatgca 900

Qy        901 ctggatagatatggaaggactgctctcatacttgctgtatgttgtggatcagcaagtata 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 ctggatagatatggaaggactgctctcatacttgctgtatgttgtggatcagcaagtata 960

Qy        961 gtcagccttctacttgagcaaaatattgatgtatcttctcaagatctatctggacagacg 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 gtcagccttctacttgagcaaaatattgatgtatcttctcaagatctatctggacagacg 1020

Qy       1021 gccagagagtatgctgtttctagtcatcatcatgtaatttgccagttactttctgactac 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 gccagagagtatgctgtttctagtcatcatcatgtaatttgccagttactttctgactac 1080

Qy       1081 aaagaaaaacagatgctaaaaatctcttctgaaaacagcaatccagaaaatgtctcaaga 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 aaagaaaaacagatgctaaaaatctcttctgaaaacagcaatccagaaaatgtctcaaga 1140

Qy       1141 accagaaataaataa 1155
              |||||||||||||||
Db       1141 accagaaataaataa 1155
```

```
RESULT   6
AAC81012
ID   AAC81012 standard; cDNA; 2000 BP.
XX
AC   AAC81012;
XX
DT   13-FEB-2001  (first entry)
XX
DE   Human B11Ag1 antigen splice isoform B11C-8 cDNA.
XX
KW   Human; breast tumour-specific antigen; cytostatic; vaccine;
KW   breast cancer; B18Ag1; B11Ag1; B15Ag1; ss.
XX
OS   Homo sapiens.
XX
PN   WO200061753-A2.
XX
PD   19-OCT-2000.
XX
PF   07-APR-2000; 2000WO-US09312.
XX
PR   09-APR-1999;   99US-0289198.
PR   28-OCT-1999;   99US-0429755.
PR   23-MAR-2000; 2000US-0534825.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Frudakis TN,  Smith JM,  Reed SG,  Misher LE,  Retter MW,  Dillon DC;
XX
DR   WPI; 2000-628403/60.
DR   P-PSDB; AAB28629.
XX
PT   An isolated polypeptide comprising an immunogenic portion of a breast
PT   tumor protein used for inhibiting the development of cancer, especially
PT   breast cancer, and monitoring cancer progression in a patient -
XX
PS   Claim 4; Page 177-178; 187pp; English.
XX
CC   The present sequence is given in a specification relating to compositions
CC   and methods for the treatment and diagnosis of breast cancer. Nucleotide
CC   sequences that are preferentially expressed in breast tumour tissue, and
CC   the polypeptides encoded by such nucleotide sequences, are used in
CC   compositions and vaccines to inhibit the development of cancer,
CC   especially breast cancer. The progression of a cancer may be monitored by
CC   carrying out detection of tumour-specific antigens at subsequent time
CC   points and comparing the results from the different time points.
CC   CD4+ and/or CD8+ T-Cells isolated from the cancer patient may be treated
CC   with tumour-specific polypeptides, polynucleotides encoding the
CC   polypeptides or antigen presenting cells expressing the polypeptides. The
CC   cells are then administered to the patient to inhibit development of
CC   cancer.
XX
SQ   Sequence 2000 BP; 698 A; 388 C; 489 G; 425 T; 0 other;

  Query Match             97.7%;  Score 1128;  DB 21;  Length 2000;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1128;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 atggtggttgaggttgattccatgccggctgcctcttctgtgaagaagccatttggtctc 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db          1 atggtggttgaggttgattccatgccggctgcctcttctgtgaagaagccatttggtctc 60

Qy         61 aggagcaagatgggcaagtggtgctgccgttgcttcccctgctgcagggagagcggcaag 120
              |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db         61 aggagcaagatgggcaagtggtgctgccgttgcttcccctgctgcagggagagcggcaag 120
```

```
Qy        121 agcaacgtgggcacttctggagaccacgacgactctgctatgaagacactcaggagcaag 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 agcaacgtgggcacttctggagaccacgacgactctgctatgaagacactcaggagcaag 180

Qy        181 atgggcaagtggtgccgccactgcttcccctgctgcaggggagtggcaagagcaacgtg 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 atgggcaagtggtgccgccactgcttcccctgctgcaggggagtggcaagagcaacgtg 240

Qy        241 ggcgcttctggagaccacgacgactctgctatgaagacactcaggaacaagatgggcaag 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ggcgcttctggagaccacgacgactctgctatgaagacactcaggaacaagatgggcaag 300

Qy        301 tggatgctgccactgcttcccctgctgcaggggagcggcaagagcaaggtgggcgcttgg 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 tggatgctgccactgcttcccctgctgcaggggagcggcaagagcaaggtgggcgcttgg 360

Qy        361 ggagactacgatgacagtgccttcatggagcccaggtaccacgtccgtggagaagatctg 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ggagactacgatgacagtgccttcatggagcccaggtaccacgtccgtggagaagatctg 420

Qy        421 gacaagctccacagagctgcctggtggggtaaagtccccagaaaggatctcatcgtcatg 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 gacaagctccacagagctgcctggtggggtaaagtccccagaaaggatctcatcgtcatg 480

Qy        481 ctcagggacactgacgtgaacaagaaggacaagcaaaagaggactgctctacatctggcc 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ctcagggacactgacgtgaacaagaaggacaagcaaaagaggactgctctacatctggcc 540

Qy        541 tctgccaatgggaattcagaagtagtaaaactcctgctggacagacgatgtcaacttaat 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 tctgccaatgggaattcagaagtagtaaaactcctgctggacagacgatgtcaacttaat 600

Qy        601 gtccttgacaacaaaaagaggacagctctgataaaggccgtacaatgccaggaagatgaa 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 gtccttgacaacaaaaagaggacagctctgataaaggccgtacaatgccaggaagatgaa 660

Qy        661 tgtgcgttaatgttgctggaacatggcactgatccaaatattccagatgagtatggaaat 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 tgtgcgttaatgttgctggaacatggcactgatccaaatattccagatgagtatggaaat 720

Qy        721 accactctgcactacgctatctataatgaagataaattaatggccaaagcactgctctta 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 accactctgcactacgctatctataatgaagataaattaatggccaaagcactgctctta 780

Qy        781 tatggtgctgatatcgaatcaaaaaacaagcatggcctcacaccactgttacttggtgta 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 tatggtgctgatatcgaatcaaaaaacaagcatggcctcacaccactgttacttggtgta 840

Qy        841 catgagcaaaaacagcaagtcgtgaaatttttaatcaagaaaaaagcgaatttaaatgca 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 catgagcaaaaacagcaagtcgtgaaatttttaatcaagaaaaaagcgaatttaaatgca 900

Qy        901 ctggatagatatggaaggactgctctcatacttgctgtatgttgtggatcagcaagtata 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 ctggatagatatggaaggactgctctcatacttgctgtatgttgtggatcagcaagtata 960

Qy        961 gtcagccttctacttgagcaaaatattgatgtatcttctcaagatctatctggacagacg 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 gtcagccttctacttgagcaaaatattgatgtatcttctcaagatctatctggacagacg 1020

Qy       1021 gccagagagtatgctgtttctagtcatcatcatgtaatttgccagttactttctgactac 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 gccagagagtatgctgtttctagtcatcatcatgtaatttgccagttactttctgactac 1080

Qy       1081 aaagaaaaacagatgctaaaaatctcttctgaaaacagcaatccagaa 1128
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 aaagaaaaacagatgctaaaaatctcttctgaaaacagcaatccagaa 1128
```

```
RESULT      7
AAB93715
ID   AAB93715 standard; cDNA; 2000 BP.
XX
AC   AAB93715;
XX
DT   04-OCT-2001  (first entry)
XX
DE   Human prostate specific cDNA sequence B305D splice variant #9.
XX
KW   Human; prostate cancer; prostate specific; diagnosis; vaccine;
KW   cytostatic; gene therapy; metastasis; ss.
XX
OS   Homo sapiens.
XX
PN   WO200151633-A2.
XX
PD   19-JUL-2001.
XX
PF   16-JAN-2001; 2001WO-US01574.
XX
PR   14-JAN-2000; 2000US-0483672.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Xu J,  Dillon DC,  Mitcham JL,  Harlocker SL,  Jiang Y,  Reed SG;
PI   Kalos MD,  Fanger GR,  Day CH,  Retter MW,  Stolk JA,  Skeiky YAW;
PI   Wang A,  Meagher MJ;
XX
DR   WPI; 2001-425873/45.
XX
PT   New polynucleotide encoding a prostate-specific protein, for
PT   diagnosing, monitoring and treating prostate cancer in a patient and
PT   for use in vaccines -
XX
PS   Claim 1; Page 347-348; 543pp; English.
XX
CC   The present invention describes polynucleotide sequences (I) which encode
CC   prostate-specific proteins (II). (I) and (II) have cytostatic activity,
CC   and can be used in vaccine production and gene therapy. (I), (II),
CC   antibodies to (II), fusion proteins comprising (II), and isolated
CC   T cells prepared using (I) or (II) are used treat cancer in a patient
CC   (I) and the antibodies are also used in the detection of cancer in a
CC   patient. The cancer that is diagnosed or treated is particularly
CC   prostate cancer. (I) and (II) can be used in vaccines. The antibodies or
CC   (I) can be used for monitoring the progression of cancer in a patient
CC   (I) and (II) can also be used to improve diagnostic and therapeutic
CC   methods for prostate cancer. They can indicate the level of metastasis
CC   as well as the prostate volume. AAB93357 to AAB93944 and AAM01113 to
CC   AAM01318 represent polynucleotide and amino acid sequences used in the
CC   exemplification of the present invention.
XX
SQ   Sequence 2000 BP; 698 A; 388 C; 489 G; 425 T; 0 other;

  Query Match             97.7%;  Score 1128;  DB 22;  Length 2000;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1128;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 atggtggttgaggttgattccatgccggctgcctcttctgtgaagaagccatttggtctc 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 atggtggttgaggttgattccatgccggctgcctcttctgtgaagaagccatttggtctc 60

Qy         61 aggagcaagatgggcaagtggtgctgccgttgcttcccctgctgcaggggagcggcaag 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 aggagcaagatgggcaagtggtgctgccgttgcttcccctgctgcaggggagcggcaag 120

Qy        121 agcaacgtgggcacttctggagaccacgacgactctgctatgaagacactcaggagcaag 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 agcaacgtgggcacttctggagaccacgacgactctgctatgaagacactcaggagcaag 180

Qy        181 atgggcaagtggtgccgccactgcttcccctgctgcaggggagtggcaagagcaacgtg 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 atgggcaagtggtgccgccactgcttcccctgctgcaggggagtggcaagagcaacgtg 240

Qy        241 ggcgcttctggagaccacgacgactctgctatgaagacactcaggaacaagatgggcaag 300
```

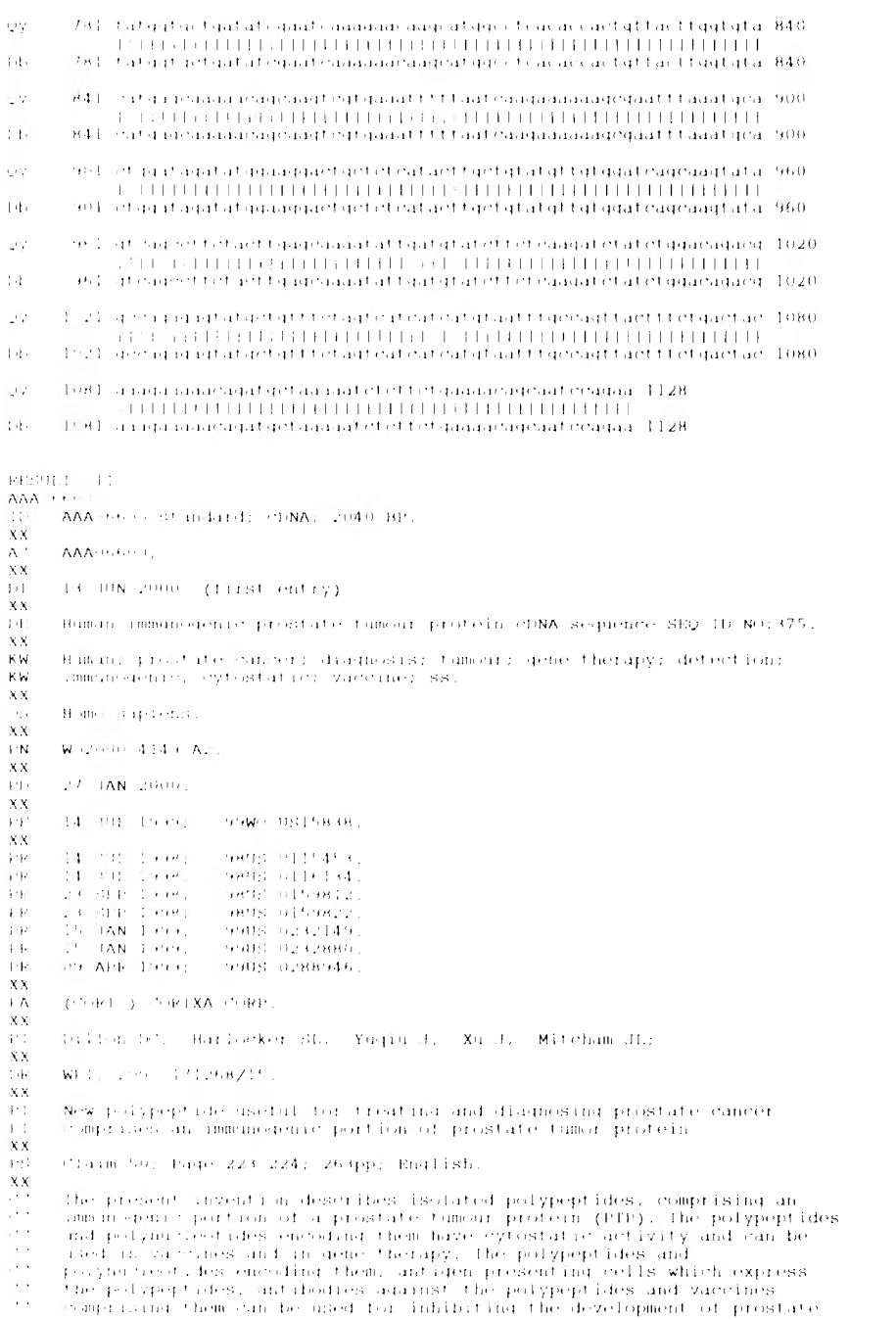
```
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 ggcgcttctggagaccacgacgactctgctatgaagacactcaggaacaagatgggcaag 300

Qy      301 tggtgctgccactgcttcccctgctgcaggggggagcggcaagagcaaggtgggcgcttgg 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 tggtgctgccactgcttcccctgctgcaggggggagcggcaagagcaaggtgggcgcttgg 360

Qy      361 ggagactacgatgacagtgccttcatggagcccaggtaccacgtccgtggagaagatctg 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      361 ggagactacgatgacagtgccttcatggagcccaggtaccacgtccgtggagaagatctg 420

Qy      421 gacaagctccacagagctgcctggtggggtaaagtccccagaaaggatctcatcgtcatg 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      421 gacaagctccacagagctgcctggtggggtaaagtccccagaaaggatctcatcgtcatg 480

Qy      481 ctcagggacactgacgtgaacaagaaggacaagcaaaagaggactgctctacatctggcc 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      481 ctcagggacactgacgtgaacaagaaggacaagcaaaagaggactgctctacatctggcc 540

Qy      541 tctgccaatgggaattcagaagtagtaaaactcctgctggacagacgatgtcaacttaat 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      541 tctgccaatgggaattcagaagtagtaaaactcctgctggacagacgatgtcaacttaat 600

Qy      601 gtccttgacaacaaaaagaggacagctctgataaaggccgtacaatgccaggaagatgaa 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      601 gtccttgacaacaaaaagaggacagctctgataaaggccgtacaatgccaggaagatgaa 660

Qy      661 tgtgcgttaatgttgctggaacatggcactgatccaaatattccagatgagtatggaaat 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      661 tgtgcgttaatgttgctggaacatggcactgatccaaatattccagatgagtatggaaat 720

Qy      721 accactctgcactacgctatctataatgaagataaattaatggccaaagcactgctctta 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      721 accactctgcactacgctatctataatgaagataaattaatggccaaagcactgctctta 780

Qy      781 tatggtgctgatatcgaatcaaaaaacaagcatggcctcacaccactgttacttggtgta 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      781 tatggtgctgatatcgaatcaaaaaacaagcatggcctcacaccactgttacttggtgta 840

Qy      841 catgagcaaaaacagcaagtcgtgaaatttttaatcaagaaaaaagcgaatttaaatgca 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      841 catgagcaaaaacagcaagtcgtgaaatttttaatcaagaaaaaagcgaatttaaatgca 900

Qy      901 ctggatagatatggaaggactgctctcatacttgctgtatgttgtggatcagcaagtata 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      901 ctggatagatatggaaggactgctctcatacttgctgtatgttgtggatcagcaagtata 960

Qy      961 gtcagccttctacttgagcaaaatattgatgtatcttctcaagatctatctggacagacg 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      961 gtcagccttctacttgagcaaaatattgatgtatcttctcaagatctatctggacagacg 1020

Qy     1021 gccagagagtatgctgtttctagtcatcatcatgtaatttgccagttactttctgactac 1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1021 gccagagagtatgctgtttctagtcatcatcatgtaatttgccagttactttctgactac 1080

Qy     1081 aaagaaaaacagatgctaaaaatctcttctgaaaacagcaatccagaa 1128
            ||||||||||||||||||||||||||||||||||||||||||||||||
Db     1081 aaagaaaaacagatgctaaaaatctcttctgaaaacagcaatccagaa 1128
```

RESULT   8
AAH85029
ID   AAH85029 standard; cDNA; 2000 BP.
XX
AC   AAH85029;
XX
DT   25-SEP-2001 (first entry)
XX
DE   Human prostate-specific cDNA sequence B305D splice variant #9.
XX
KW   Human; prostate cancer; therapy; diagnosis; cat eye syndrome;
KW   chromosome 22q11.2; prostate specific protein, chromosome 1,
KW   prostate specific antigen; PSA; ss.
XX
OS   Homo sapiens.
XX
PN   WO200134802-A2.
XX
PD   17-MAY-2001.
XX
PF   03-NOV-2000; 2000WO-US30094.
XX
PR   12-NOV-1999;   99US-0439313.
PR   18-NOV-1999;   99US-0443686.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG;
PI   Kalos MD, Retter MW, Stolk JA, Day CH, Skeiky YAW, Wang A;
XX
DR   WPI; 2001-408785/32.
XX
PT   Isolated polypeptide comprising at least an immunogenic portion of a
PT   prostate-specific protein, useful in the diagnosis and therapy of
PT   prostate cancer -
XX
PS   Claim 31; Page 247-248; 325pp; English.
XX
CC   The present invention describes an isolated polypeptide (P1) comprising
CC   at least an immunogenic portion of a prostate specific protein, or its
CC   variant. Also described are polynucleotides (N1; encoding (P1). (P1) and
CC   (N1) have cytostatic activity and can be used in vaccine production.
CC   The polypeptides, nucleic acids and antibodies from the present
CC   invention are useful in the diagnosis and therapy of prostate cancer.
CC   Prostate specific genes P704P, P712P, P774P, P775P and B305D are located
CC   in a genomic region on chromosome 22q11.2 known as the Cat Eye Syndrome
CC   region. Prostate specific antigen (PSA) P501S was located on
CC   chromosome 1. AAH84671 to AAH85143 and AAG99000 to AAG99077 represent
CC   polynucleotide and polypeptide sequences used in the exemplification
CC   of the present invention.
XX
SQ   Sequence 2000 BP; 698 A; 388 C; 489 G; 425 T; 0 other;

```
  Query Match             97.7%;  Score 1128;  DB 22;  Length 2000;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1128;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 atggtggttgaggttgattccatgccggctgcctcttctgtgaagaagccatttggtctc 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 atggtggttgaggttgattccatgccggctgcctcttctgtgaagaagccatttggtctc 60

Qy       61 aggagcaagatgggcaagtggtgctgccgttgcttcccctgctgcagggagagagcggcaag 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 aggagcaagatgggcaagtggtgctgccgttgcttcccctgctgcagggagagagcggcaag 120

Qy      121 agcaacgtgggcacttctggagaccacgacgactctgctatgaagacactcaggagcaag 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 agcaacgtgggcacttctggagaccacgacgactctgctatgaagacactcaggagcaag 180

Qy      181 atgggcaagtggtgccgccactgcttcccctgctgcaggggggagtggcaagagcaacgtg 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 atgggcaagtggtgccgccactgcttcccctgctgcaggggggagtggcaagagcaacgtg 240

Qy      241 ggcgcttctggagaccacgacgactctgctatgaagacactcaggaacaagatgggcaag 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 ggcgcttctggagaccacgacgactctgctatgaagacactcaggaacaagatgggcaag 300

Qy      301 tggtgctgccactgcttcccctgctgcaggggggagcggcaagagcaaggtgggcgcttgg 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 tggtgctgccactgcttcccctgctgcaggggggagcggcaagagcaaggtgggcgcttgg 360

Qy      361 ggagactacgatgacagtgccttcatggagcccaggtaccacgtccgtggagaagatctg 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      361 ggagactacgatgacagtgccttcatggagcccaggtaccacgtccgtggagaagatctg 420
```

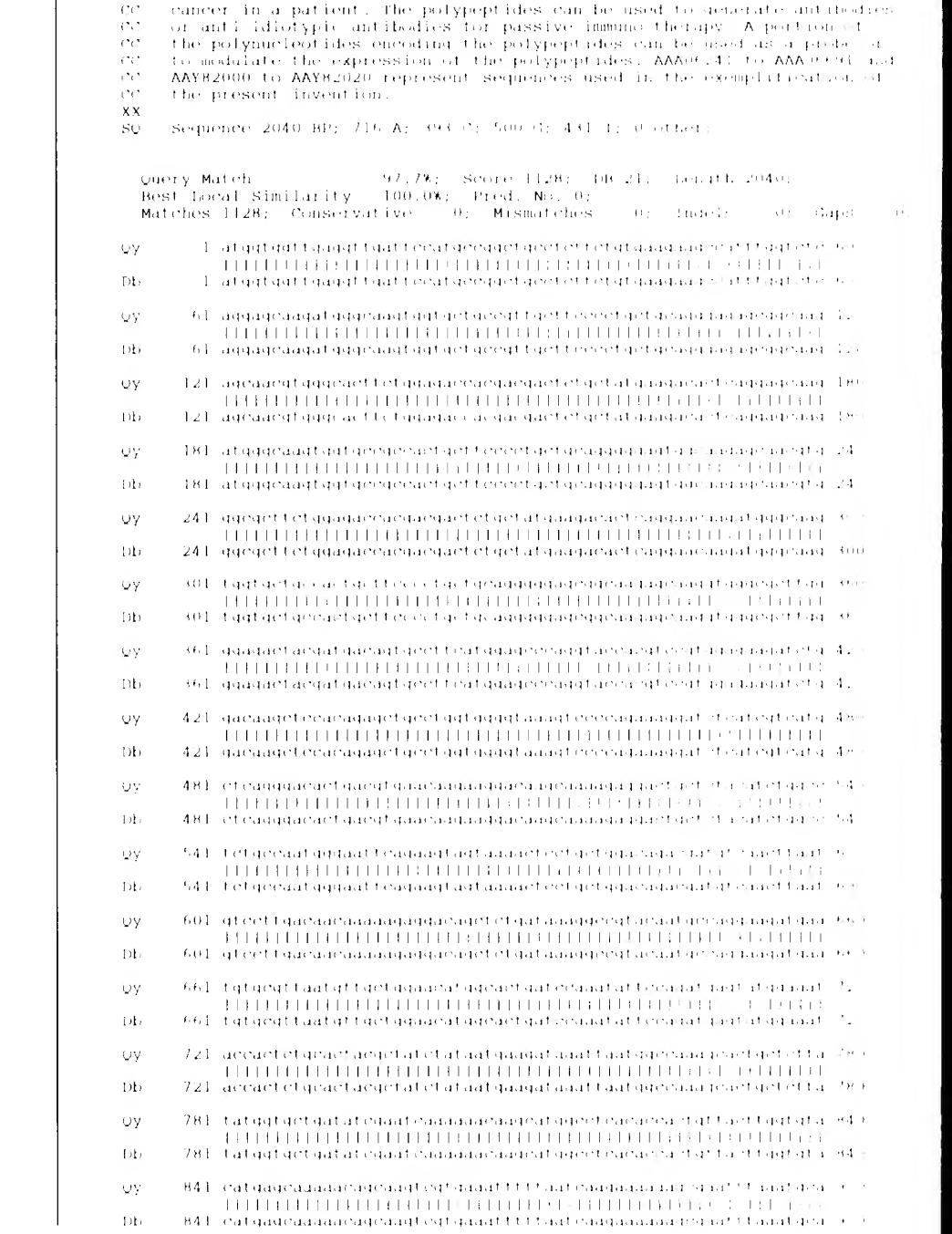
```
Qy      421 qacaaqctccacaaqactqccttqqtqqqqtaaaqtcccaqaaaaqqatctcatcqtcatq 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      421 qacaaqctccacaqaactqccttqqtqqqqtaaaqtcccaqaaaaqqatctcatcqtcatq 480

Qy      481 ctccaqqaacactqacqtqaacaaqaaqqacaaqcaaaaqaqactqctctacatctqqcc 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      481 ctccaqqaacactqacqtqaacaaqaaqqacaaqcaaaaqaqactqctctacatctqqcc 540

Qy      541 tctqccaatqqqaattcaqaaqtaqtaaaactcctqctqqacaqacqatqtcaacttaat 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      541 tctqccaatqqqaattcaqaaqtaqtaaaactcctqctqqacaqacqatqtcaacttaat 600

Qy      601 qtccttqacaacaaaaaqaqqacaqctctqataaaqqccqtacaatqccaqqaaqatqaa 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      601 qtccttqacaacaaaaaqaqqacaqctctqataaaqqccqtacaatqccaqqaaqatqaa 660

Qy      661 tqtqcqttaatqttqctqqaacatqqcactqatccaaatattccaqatqaqtatqqaaat 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      661 tqtqcqttaatqttqctqqaacatqqcactqatccaaatattccaqatqaqtatqqaaat 720

Qy      721 accactctqcactacqctatctataatqaaqataaattaatqqccaaaqcactqctctta 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      721 accactctqcactacqctatctataatqaaqataaattaatqqccaaaqcactqctctta 780

Qy      781 tatqqtqctqatatcqaatcaaaaaacaaqcatqqcctcacaccactqttacttqqtqta 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      781 tatqqtqctqatatcqaatcaaaaaacaaqcatqqcctcacaccactqttacttqqtqta 840

Qy      841 catqqaccaaaacaqcaaqtcqtqaaattttaatcaaqaaaaaqcqaatttaaatqca 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      841 catqqaccaaaacaqcaaqtcqtqaaattttaatcaaqaaaaaqcqaatttaaatqca 900

Qy      901 ctqqatqatatqaaaqqactqctctcatacttqctqtatqttqtqqatcaqcaaqtata 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      901 ctqqatqatatqaaaqqactqctctcatacttqctqtatqttqtqqatcaqcaaqtata 960

Qy      961 qtcaqcttctacttqaqcaaaatattqatqtqtatcttctcaaqatctatctqqacaqaq 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      961 qtcaqcttctacttqaqcaaaatattqatqtqtatcttctcaaqatctatctqqacaqaq 1020

Qy     1021 qccaqaqqatatqctqtttctaqtcatcatcatqtaattqccaqttacttttctqactac 1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1021 qccaqaqqatatqctqtttctaqtcatcatcatqtaattqccaqttacttttctqactac 1080

Qy     1081 aaaaaaaaacaqatqctaaaaatctcttctqaaaacaqcaatccaqaa 1128
            |||||||||||||||||||||||||||||||||||||||||||||||||
Db     1081 aaaaaaaaacaqatqctaaaaatctcttctqaaaacaqcaatccaqaa 1128


RESULT 9
AAB02780
ID   AAB02780 standard; cDNA; 2000 BP.
XX
AC   AAB02780;
XX
DT   14-JUN-2001 (first entry)
XX
DE   Prostate tumour antigen determined cDNA splice variant of B305D #9.
XX
KW   Human; prostate tumour antigen; prostate tumour; therapy; diagnosis;
KW   prostate cancer; immunogenic; cytostatic; vaccine; ss.
XX
OS   Homo sapiens.
XX
PN   WO200125272-A2.
XX
PD   12-APR-2001.
XX
PF   04-OCT-2000; 2000WO-US27464.
XX
PR   04-OCT-1999; 99US-015745.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Xu J, Skeiky YAW, Reed SG, Choever MA;
XX
DR   WPI; 2001-245062/25.
DR   P-PSDB; AAB74816.
XX
PT   Prostate specific protein and its encoding polynucleotide, useful for
PT   the treatment and diagnosis of prostate cancer
XX
PS   Claim 50; Page 232; 276pp; English.
XX
CC   The present invention describes an isolated polypeptide (I) comprising
CC   at least an immunogenic portion of a prostate tumour antigen protein or
CC   its variant. (I) have cytostatic activity and can be used in vaccine
CC   production. (I), prostate tumour antigen polynucleotides, an antigen
CC   presenting cell (APC e.g. a dendritic cell) that expresses (I), and a
CC   pharmaceutical composition containing (I) are useful for inhibiting the
CC   development of cancer in a patient. Antibodies specific for prostate
CC   specific proteins and oligonucleotides that hybridise to a
CC   polynucleotide that encodes a prostate specific protein are useful
CC   for detecting the presence or absence of a cancer or monitoring the
CC   progression the progression of a cancer, especially prostate cancer.
CC   AAH02422 to AAH02872, AAB74798 to AAB74821 and AAB74830 are sequences
CC   used in the exemplification of the present invention.
XX
SQ   Sequence 2000 BP; 698 A; 388 C; 489 G; 425 T; 0 other


  Query Match             97.7%;  Score 1128;  DB 22;  Length 2000;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1128;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 atqqtqqttqaqqttqattccatqccqqctqcctcttctqtqaaaaaqccatttqqtctc 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 atqqtqqttqaqqttqattccatqccqqctqcctcttctqtqaaaaaqccatttqqtctc 60

Qy       61 aqqaqcaaqatqqqcaaqtqqtqctqccqttqcttcccctqctqccaqqqaaaaqccqcaaq 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 aqqaqcaaqatqqqcaaqtqqtqctqccqttqcttcccctqctqccaqqqaaaaqccqcaaq 120

Qy      121 aqcaaccqtqqqcacttctqqaqaccacqacqactctqctatqaaqaactcaqqaqcaaq 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 aqcaaccqtqqqcacttctqqaqaccacqacqactctqctatqaaqaactcaqqaqcaaq 180

Qy      181 atqqqcaaqtqqtqccqccactqcttcccctqctqccaqqqqqaqtqqcaaqaqcaaqtq 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 atqqqcaaqtqqtqccqccactqcttcccctqctqccaqqqqqaqtqqcaaqaqcaaqtq 240

Qy      241 qqcqcttctqqaqaccacqacqactctqctatqaaqaactcaqqaqcaaqatqqqcaaq 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 qqcqcttctqqaqaccacqacqactctqctatqaaqaactcaqqaqcaaqatqqqcaaq 300

Qy      301 tqqtqctqccactqcttcccctqctqccaqqqqqaqtqqcaaqaqcaaqtqqcqcttqa 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 tqqtqctqccactqcttcccctqctqccaqqqqqaqtqqcaaqaqcaaqtqqcqcttqa 360

Qy      361 qqaqactacqatqacaqtqcttcatqaaqccaqqtaccacqctccqtqaaaaatctq 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      361 qqaqactacqatqacaqtqcttcatqaaqccaqqtaccacqctccqtqaaaaatctq 420

Qy      421 qacaaqctccacaqaactqccttqqtqqqqtaaaqtcccaqaaaaqqatctcatcqtcatq 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      421 qacaaqctccacaqaactqccttqqtqqqqtaaaqtcccaqaaaaqqatctcatcqtcatq 480

Qy      481 ctccaqqaacactqacqtqaacaaqaaqqacaaqcaaaaqaqactqctctacatctqqcc 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      481 ctccaqqaacactqacqtqaacaaqaaqqacaaqcaaaaqaqactqctctacatctqqcc 540

Qy      541 tctqccaatqqqaattcaqaaqtaqtaaaactcctqctqqacaqacqatqtcaacttaat 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      541 tctqccaatqqqaattcaqaaqtaqtaaaactcctqctqqacaqacqatqtcaacttaat 600

Qy      601 qtccttqacaacaaaaaqaqqacaqctctqataaaqqccqtacaatqccaqqaaqatqaa 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      601 qtccttqacaacaaaaaqaqqacaqctctqataaaqqccqtacaatqccaqqaaqatqaa 660
```

```
Db    601 gtccttgacaacaaaaagaggacagctctgataaagggccgtacaatgccagggaaagatgaa 660

Qy    661 tgtgcgttaatgttgctgggaacatggcactgatccaaatattccagatgagtatgaaaat 720

Db    661 tgtgcgttaatgttgctgggaacatggcactgatccaaatattccagatgagtatgaaaat 720

Qy    721 accactctgcactacgctatctataatgaagatgaaattaatgggccaaagcactgctctta 780

Db    721 accactctgcactacgctatctataatgaagatgaaattaatgggccaaagcactgctctta 780

Qy    781 tatggtgctgatatcgaatcaaaaaacaagcatggcctcacaccactgttacttggtgta 840

Db    781 tatggtgctgatatcgaatcaaaaaacaagcatggcctcacaccactgttacttggtgta 840

Qy    841 catgagcaaaaacagcaagtcgtgaaattttaatcaagaaaaaagcgaatttaaatgca 900

Db    841 catgagcaaaaacagcaagtcgtgaaattttaatcaagaaaaaagcgaatttaaatgca 900

Qy    901 ctggatagatatggaaggactgctctcatacttgctgtatgttgtggatcagcaagtata 960

Db    901 ctggatagatatggaaggactgctctcatacttgctgtatgttgtggatcagcaagtata 960

Qy    961 gtcagccttctacttgagcaaaatattgatgtatcttctcaagatctatctgacagacg 1020

Db    961 gtcagccttctacttgagcaaaatattgatgtatcttctcaagatctatctggacagacg 1020

Qy   1021 gccagagaagtatgctgtttctagtcatcatcatgtaatttgccagttacttttctgactac 1080

Db   1021 gccagagaagtatgctgtttctagtcatcatcatgtaatttgccagttacttttctgactac 1080

Qy   1081 aaagaaaaacagatgctaaaaatctcttctgaaaacagcaatccagaa 1128

Db   1081 aaagaaaaacagatgctaaaaatctcttctgaaaacagcaatccagaa 1128


RESULT 10
AAC81013
ID   AAC81013 standard; cDNA; 2040 BP.
XX
AC   AAC81013;
XX
DT   13-FEB-2001 (first entry)
XX
DE   Human B11Ag1 antigen splice isoform B11C-9-16 cDNA.
XX
KW   Human; breast tumour-specific antigen; cytostatic; vaccine;
KW   breast cancer; B18Ag1; B11Ag1; B15Ag1; ss.
XX
OS   Homo sapiens.
XX
PN   WO200061753-A2.
XX
PD   19-OCT-2000.
XX
PF   07-APR-2000; 2000WO-US09312.
XX
PR   09-APR-1999;  99US-0289198.
PR   28-OCT-1999;  99US-0429755.
PR   23-MAR-2000;  2000US-0534825.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Prudakis IN, Smith JM, Reed SG, Misher LE, Retter MW, Dillon DC;
XX
DR   WP1; 2000-628403/60.
DR   P-PSDB; AAB28630.
XX
PT   An isolated polypeptide comprising an immunogenic portion of a breast
PT   tumor protein used for inhibiting the development of cancer, especially
PT   breast cancer, and monitoring cancer progression in a patient -
XX
PS   Claim 4; Page 178; 187pp; English.

XX
CC   The present sequence is given in a specification relating to compositions
CC   and methods for the treatment and diagnosis of breast cancer. Nucleotide
CC   sequences that are preferentially expressed in breast tumour tissue, and
CC   the polypeptides encoded by such nucleotide sequences, are used in
CC   compositions and vaccines to inhibit the development of cancer,
CC   especially breast cancer. The progression of a cancer may be monitored by
CC   carrying out detection of tumour-specific antigens at subsequent time
CC   points and comparing the results from the different time points.
CC   CD4+ and/or CD8+ T cells isolated from the cancer patient may be treated
CC   with tumour-specific polypeptides, polynucleotides encoding the
CC   polypeptides or antigen presenting cells expressing the polypeptides. The
CC   cells are then administered to the patient to inhibit development of
CC   cancer.
XX
SQ   Sequence 2040 BP; 716 A; 392 C; 500 G; 432 T; 0 other;


  Query Match             97.7%;  Score 1128;  DB 21;  Length 2040;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1128;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 atggtggttgaggttgattccatgccggctgcctcttctgtgaagaaagccatttggtctc 60

Db      1 atggtggttgaggttgattccatgccggctgcctcttctgtgaagaaagccatttggtctc 60

Qy     61 aggagcaagatgtggcaagtggtgctgccgttgcttccccctgctgccaggcaagaagcagcaag 120

Db     61 aggagcaagatgtggcaagtggtgctgccgttgcttccccctgctgccaggcaagaagcagcaag 120

Qy    121 agcaacgtgggcacttctgggagaccacgacgactctgctatgaagacactcaggaggcaag 180

Db    121 agcaacgtgggcacttctgggagaccacgacgactctgctatgaagacactcaggaggcaag 180

Qy    181 atgggcaagtggtgccgccactgcttccccctgctgcaggcggcagtggcaagaagcaacgtg 240

Db    181 atgggcaagtggtgccgccactgcttccccctgctgcaggcggcagtggcaagaagcaacgtg 240

Qy    241 ggcgcttctgggagaccacgacgactctgctatgaagacactcaggaagcaagatgcggcaag 300

Db    241 ggcgcttctgggagaccacgacgactctgctatgaagacactcaggaagcaagatgcggcaag 300

Qy    301 tggtgctgccactgcttccccctgctgcaggcggcagtggcaagaagcaagatgcggcacttgg 360

Db    301 tggtgctgccactgcttccccctgctgcaggcggcagtggcaagaagcaagatgcggcacttgg 360

Qy    361 ggagactacgatgacagtgccttcatggaggccaggtaccacgtccgtgcagaaaatatctg 420

Db    361 ggagactacgatgacagtgccttcatggaggccaggtaccacgtccgtgcagaaaatatctg 420

Qy    421 gacaagctccacagagctgcctggtgggtaaagtccccagaaaggatctcatcgtcatg 480

Db    421 gacaagctccacagagctgcctggtgggtaaagtccccagaaaggatctcatcgtcatg 480

Qy    481 ctcagggacactgacgtgaacaagaaggacaagcaaaagaggactgctctacatctggcc 540

Db    481 ctcagggacactgacgtgaacaagaaggacaagcaaaagaggactgctctacatctggcc 540

Qy    541 tctgccaatgggaattcagaagtagtaaaactcctgctgcagacagacgatgtcaacttaat 600

Db    541 tctgccaatgggaattcagaagtagtaaaactcctgctgcagacagacgatgtcaacttaat 600

Qy    601 gtccttgacaacaaaaagaggacagctctgataaagggccgtacaatgccagggaaagatgaa 660

Db    601 gtccttgacaacaaaaagaggacagctctgataaagggccgtacaatgccagggaaagatgaa 660

Qy    661 tgtgcgttaatgttgctgggaacatggcactgatccaaatattccagatgagtatgaaaat 720

Db    661 tgtgcgttaatgttgctgggaacatggcactgatccaaatattccagatgagtatgaaaat 720

Qy    721 accactctgcactacgctatctataatgaagatgaaattaatgggccaaagcactgctctta 780

Db    721 accactctgcactacgctatctataatgaagatgaaattaatgggccaaagcactgctctta 780
```



RESULT 11
AAA96661
ID AAA96661 standard; cDNA; 2040 BP.
XX
AC AAA96661;
XX
DT 14-JUN-2000 (first entry)
XX
DE Human immunogenic prostate tumour protein cDNA sequence SEQ ID NO:475.
XX
KW Human; prostate cancer; diagnosis; tumour; gene therapy; detection;
KW immunogenic; cytostatic; vaccine; ss.
XX
OS Homo sapiens.
XX
PN WO200044894-A2.
XX
PD 27-JAN-2000.
XX
PF 14-JUL-1999; 99WO-US015848.
XX
PR 24-JUL-1998; 98US-0117454.
PR 24-JUL-1998; 98US-0117464.
PR 29-SEP-1998; 98US-0102052.
PR 29-SEP-1998; 98US-0102222.
PR 25-JAN-1999; 99US-0237149.
PR 27-JAN-1999; 99US-0237887.
PR 09-APR-1999; 99US-0288946.
XX
PA (CORI) CORIXA CORP.
XX
PI Dillon DC; Hartzoker SL; Yuqin J; Xu J; Mitcham JL;
XX
DR WPI; 2000-171268/15.
XX
PT New polypeptide useful for treating and diagnosing prostate cancer
PT comprises an immunogenic portion of prostate tumor protein
XX
PS Claim 50; Page 223-224; 264pp; English.
XX
CC The present invention describes isolated polypeptides, comprising an
CC immunogenic portion of a prostate tumour protein (PTP). The polypeptides
CC and polynucleotides encoding them have cytostatic activity and can be
CC used in vaccines and in gene therapy. The polypeptides and
CC polynucleotides encoding them, antigen presenting cells which express
CC the polypeptides, antibodies against the polypeptides and vaccines
CC comprising them can be used for inhibiting the development of prostate
CC cancer in a patient. The polypeptides can be used to generate antibodies
CC or anti-idiotypic antibodies for passive immune therapy. A portion of
CC the polynucleotides encoding the polypeptides can be used as a probe or
CC to modulate the expression of the polypeptides. AAA96641 to AAA96661 and
CC AAY82000 to AAY82020 represent sequences used in the exemplification of
CC the present invention.
XX
SQ Sequence 2040 BP; 716 A; 393 C; 500 G; 431 T; 0 other;

Query Match 97.7%; Score 1128; DB 21; Length 2040;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



```
Qy     901 ctggatagatatgqaaqgactgctctcatacttgctgtatgttgtggatcagcaagtata 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 ctggatagatatgqaaggactgctctcatacttgctgtatgttgtggatcagcaagtata 960

Qy     961 gtcagccttctacttgagcaaaatattgatgtatcttctcaagatctatctggacagacg 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 gtcagccttctacttgagcaaaatattgatgtatcttctcaagatctatctggacagacg 1020

Qy    1021 gccagaggagtatgctgtttctagtcatcatcatgtaatttgccagttacttttctgactac 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 gccagaggagtatgctgtttctagtcatcatcatgtaatttgccagttacttttctgactac 1080

Qy    1081 aaagaaaaaacagatgctaaaaatctcttctgaaaacagcaatccagaa 1128
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 aaagaaaaaacagatgctaaaaatctcttctgaaaacagcaatccagaa 1128


RESULT 12
AAH93716
ID   AAH93716 standard; cDNA; 2040 BP
XX
AC   AAH93716;
XX
DT   04-OCT-2001 (first entry)
XX
DE   Human prostate-specific cDNA sequence B305D splice variant #10.
XX
KW   Human; prostate cancer; prostate-specific; diagnosis; vaccine;
KW   cytostatic; gene therapy; metastasis; ss.
XX
OS   Homo sapiens.
XX
PN   WO200151633-A2.
XX
PD   19-JUL-2001.
XX
PF   16-JAN-2001; 2001WO-US01574.
XX
PR   14-JAN-2000; 2000US-0483672.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG;
PI   Kalos MD, Panzer GR, Day CP, Netter MW, Shalik SA, Skelly YAW,
PI   Wang A, Meagher MJ;
XX
DR   WPI; 2001-425874/45.
XX
PT   New polynucleotide encoding a prostate-specific protein, for
PT   diagnosing, monitoring and treating prostate cancer in a patient and
PT   for use in vaccines -
XX
PS   Claim 1; Page 448; 543pp; English.
XX
CC   The present invention describes polynucleotide sequences (I) which encode
CC   prostate-specific proteins (II). (I) and (II) have cytostatic activity,
CC   and can be used in vaccine production and gene therapy. (I), (II),
CC   antibodies to (II), fusion proteins comprising (II), and isolated
CC   T cells prepared using (I) or (II) are used treat cancer in a patient.
CC   (I) and the antibodies are also used in the detection of cancer in a
CC   patient. The cancer that is diagnosed or treated is particularly
CC   prostate cancer. (I) and (II) can be used in vaccines. The antibodies or
CC   (I) can be used for monitoring the progression of cancer in a patient.
CC   (I) and (II) can also be used to improve diagnostic and therapeutic
CC   methods for prostate cancer. They can indicate the level of metastasis
CC   as well as the prostate volume. AAH93357 to AAH93944 and AAM01115 to
CC   AAM01318 represent polynucleotide and amino acid sequences used in the
CC   exemplification of the present invention.
XX
SQ   Sequence 2040 BP; 716 A; 392 C; 580 G; 432 T; 0 other;
```

```
Query Match             97.7%;  Score 1128;  DB 22;  Length 2040;
Best Local Similarity   100.0%;  Pred. No. 0;
Matches 1128;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 atggtggttgaggttgattccatgccggctgcctcttctgtgaagaagccatttggtctc 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 atggtggttgaggttgattccatgccggctgcctcttctgtgaagaagccatttggtctc 60

Qy      61 aggagcaagatgggcaagtggtgctgccgttgcttccccctgctgcaggggagaccggcaag 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 aggagcaagatgggcaagtggtgctgccgttgcttccccctgctgcaggggagaccggcaag 120

Qy     121 aggcaacgtgggcacttctgggagaccacgagcgactctgctatgaagacactcaggagcaag 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 aggcaacgtgggcacttctgggagaccacgagcgactctgctatgaagacactcaggagcaag 180

Qy     181 atgggcaagtggtgccgccactgcttccccctgctgcaggggggagtggccaagagccaacgtg 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 atgggcaagtggtgccgccactgcttccccctgctgcaggggggagtggccaagagccaacgtg 240

Qy     241 ggcgcttctgggagaccacgacgactctgctatgaagacactcaggagacaagatgggcaag 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 ggcgcttctgggagaccacgacgactctgctatgaagacactcaggagacaagatgggcaag 300

Qy     301 tggtgctgccgctgcttccccctgctgcaggggggagcgggagaagccaagggtgggcgcttgg 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 tggtgctgccgctgcttccccctgctgcaggggggagcgggagaagccaagggtgggcgcttgg 360

Qy     361 gggagctacgatgacgagtgccttcatgggagccgaggtaccacgtccgtgggagaagatctg 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 gggagctacgatgacgagtgccttcatgggagccgaggtaccacgtccgtgggagaagatctg 420

Qy     421 gacaagctccacagagctgccttggtgggtgaaagtcccgagaaagggatctcatcgtcatg 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 gacaagctccacagagctgccttggtgggtgaaagtcccgagaaagggatctcatcgtcatg 480

Qy     481 ctcagggacactgacgtgaacaagaagaggacaagcaaaagaggagctgctctacatctgggc 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 ctcagggacactgacgtgaacaagaagaggacaagcaaaagaggagctgctctacatctgggc 540

Qy     541 tctgccaatgggaattcagaagtagtaaaactcctgctgagacagacgatgtcaacttaat 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 tctgccaatgggaattcagaagtagtaaaactcctgctgagacagacgatgtcaacttaat 600

Qy     601 gtccttgacaacaaaaaagaggacagctctgatgaaggccgtacaatgccagggaatgaaa 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 gtccttgacaacaaaaaagaggacagctctgatgaaggccgtacaatgccagggaatgaaa 660

Qy     661 tgtgcgttaatgttgctgaaacatggcactgatccaaatattccagatgagtatgaaaat 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 tgtgcgttaatgttgctgaaacatggcactgatccaaatattccagatgagtatgaaaat 720

Qy     721 accactctgcactacgctatctataatgaagatgaaattaatgccgaaagcactgctctta 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 accactctgcactacgctatctataatgaagatgaaattaatgccgaaagcactgctctta 780

Qy     781 tatggtgctgatatcgaatcaaaaaacaagcatggcctcacaccactgttacttgtgtta 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 tatggtgctgatatcgaatcaaaaaacaagcatggcctcacaccactgttacttgtgtta 840

Qy     841 catgaagcaaaaacagcaagtcgtgaaatttttaatcaagaaaaaaagcgaatttaaatgca 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 catgaagcaaaaacagcaagtcgtgaaatttttaatcaagaaaaaaagcgaatttaaatgca 900

Qy     901 ctggatagatatgqaaqgactgctctcatacttgctgtatgttgtggatcagcaagtata 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 ctggatagatatgqaaggactgctctcatacttgctgtatgttgtggatcagcaagtata 960

Qy     961 gtcagccttctacttgagcaaaatattgatgtatcttctcaagatctatctggacagacg 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 gtcagccttctacttgagcaaaatattgatgtatcttctcaagatctatctggacagacg 1020

Qy    1021 gccagaggagtatgctgtttctagtcatcatcatgtaatttgccagttacttttctgactac 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 gccagaggagtatgctgtttctagtcatcatcatgtaatttgccagttacttttctgactac 1080
```

```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 gccagagagtatgctgtttctagtcatcatcatgtaatttgccagttactttctgactac 1080

Qy       1081 aaagaaaaacagatgctaaaaatctcttctgaaaacagcaatccagaa 1128
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 aaagaaaaacagatgctaaaaatctcttctgaaaacagcaatccagaa 1128


RESULT 13
AAH85030
ID   AAH85030 standard; cDNA; 2040 BP.
XX
AC   AAH85030;
XX
DT   25-SEP-2001  (first entry)
XX
DE   Human prostate specific cDNA sequence B305D splice variant #10.
XX
KW   Human; prostate cancer; therapy; diagnosis; cat eye syndrome;
KW   chromosome 22q11.2; prostate specific protein; chromosome 1;
KW   prostate specific antigen, PSA; ss.
XX
OS   Homo sapiens.
XX
PN   WO200134802-A2.
XX
PD   17-MAY-2001.
XX
PF   09-NOV-2000; 2000WO-US30904.
XX
PR   12-NOV-1999;   99US-0439313.
PR   18-NOV-1999;   99US-0443686.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Xu J,  Dillon DC,  Mitcham JL,  Harlocker SL,  Jiang Y,  Reed SG;
PI   Kalos MD,  Retter MW,  Stolk JA,  Day CH,  Skeiky YAW,  Wang A;
XX
DR   WPI; 2001-308785/32.
XX
PT   Isolated polypeptide comprising at least an immunogenic portion of a
PT   prostate specific protein, useful in the diagnosis and therapy of
PT   prostate cancer
XX
PS   Claim 31; Page 248; 325pp; English.
XX
CC   The present invention describes an isolated polypeptide (PI) comprising
CC   at least an immunogenic portion of a prostate specific protein, or its
CC   variant. Also described are polynucleotides (NI) encoding (PI). (PI) and
CC   (NI) have cytostatic activity and can be used in vaccine production.
CC   The polypeptides, nucleic acids and antibodies from the present
CC   invention are useful in the diagnosis and therapy of prostate cancer.
CC   Prostate specific genes P704P, P712P, P774P, P775P and B305D are located
CC   in a genomic region on chromosome 22q11.2 known as the Cat Eye Syndrome
CC   region. Prostate specific antigen (PSA) P501S was located on
CC   chromosome 1. AAH84671 to AAH85143 and AAG99000 to AAG99077 represent
CC   polynucleotide and polypeptide sequences used in the exemplification
CC   of the present invention.
XX
SQ   Sequence 2040 BP; 726 A; 392 C; 500 G; 432 T; 0 other;

  Query Match             97.7%;  Score 1128;  DB 22;  Length 2040;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1128;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 atggtgatttgaggttgattccatgccggctgccctcttctgtgaagaagccatttggtctc 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 atggtgatttgaggttgattccatgccggctgccctcttctgtgaagaagccatttggtctc 60

Qy         61 aggggcaaatgggcaagtgatgctgccgttgcttcccctgctgcaggggagagcggcaag 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 aggggcaaatgggcaagtgatgctgccgttgcttcccctgctgcaggggagagcggcaag 120

Qy        121 agcaacgtgggcacttctggagaccacgacgactctgctatgaaaacatcaagagcaag 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 agcaacgtgggcacttctggagaccacgacgactctgctatgaaaacatcaagagcaag 180

Qy        181 atgggcaagtgatgccgccactgcttcccctgctgcaggggaatgacaaagcaacgta 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 atgggcaagtgatgccgccactgcttcccctgctgcaggggaatgacaaagcaacgta 240

Qy        241 gggcgcttctggagaccacgacgactctgctatgaaagacactcagaacagaatggcaag 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 gggcgcttctggagaccacgacgactctgctatgaaagacactcagaacagaatggcaag 300

Qy        301 tggtgctgccactgcttcccctgctgcaggggagacggcaagagcaagatgagcgcttag 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 tggtgctgccactgcttcccctgctgcaggggagacggcaagagcaagatgagcgcttag 360

Qy        361 ggagactacgatgacagtgccttcatggagcccaggtaccacgtcccatgaaaaagatcta 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ggagactacgatgacagtgccttcatggagcccaggtaccacgtcccatgaaaaagatcta 420

Qy        421 gacaagctccacagagctgcctggtggggtaaagtccccagaaaggatctcatcgtcata 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 gacaagctccacagagctgcctggtggggtaaagtccccagaaaggatctcatcgtcata 480

Qy        481 ctcaggacactgacgtgaacaagaaggacaagcaaaaaggactgctctatctgacc 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ctcaggacactgacgtgaacaagaaggacaagcaaaaaggactgctctatctgacc 540

Qy        541 tctgccaatgggaattcagaagtagtaaaactcctgctggacagacgatgtcaacttaat 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 tctgccaatgggaattcagaagtagtaaaactcctgctggacagacgatgtcaacttaat 600

Qy        601 gtccttgacaacaaaaagaggacagctctgataaaggccgtacaatgccagaaagatgaa 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 gtccttgacaacaaaaagaggacagctctgataaaggccgtacaatgccagaaagatgaa 660

Qy        661 tgtgcgttaatgttgctggaacatggcactgatccaaatattccagatgagtatggaaat 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 tgtgcgttaatgttgctggaacatggcactgatccaaatattccagatgagtatggaaat 720

Qy        721 accactctgcactacgctatctataatgaagataaattaatggccaaagcactgctctta 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 accactctgcactacgctatctataatgaagataaattaatggccaaagcactgctctta 780

Qy        781 tatggtgctgatatcgaatcaaaaaacaagcatggcctcacacccactgttacttgatgta 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 tatggtgctgatatcgaatcaaaaaacaagcatggcctcacacccactgttacttgatgta 840

Qy        841 catgagcaaaaacagcaagtcgtgaaatttttaatcaagaaaaaagcgaatttaaatgca 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 catgagcaaaaacagcaagtcgtgaaatttttaatcaagaaaaaagcgaatttaaatgca 900

Qy        901 ctggatagatatggaaggactgctctcatacttgctgtatgttgtggatcagcaagtata 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 ctggatagatatggaaggactgctctcatacttgctgtatgttgtggatcagcaagtata 960

Qy        961 gtcagccttctacttgagcaaaatattgatgtatcttctcaagatctatctggacagacg 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 gtcagccttctacttgagcaaaatattgatgtatcttctcaagatctatctggacagacg 1020

Qy       1021 gccagagagtatgctgtttctagtcatcatcatgtaatttgccagttactttctgactac 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 gccagagagtatgctgtttctagtcatcatcatgtaatttgccagttactttctgactac 1080

Qy       1081 aaagaaaaacagatgctaaaaatctcttctgaaaacagcaatccagaa 1128
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 aaagaaaaacagatgctaaaaatctcttctgaaaacagcaatccagaa 1128


RESULT 14
AAH02781
ID   AAH02781 standard; cDNA; 2040 BP.
```

```
XX
AC   AAH02781;
XX
DT   14-JUN-2001  (first entry)
XX
DE   Prostate tumour antigen determined cDNA splice variant of B305D #10.
XX
KW   Human; prostate tumour antigen; prostate tumour; therapy; diagnosis;
KW   prostate cancer; immunogenic; cytostatic; vaccine; ss.
XX
OS   Homo sapiens.
XX
PN   WO200125272-A2.
XX
PD   12-APR-2001.
XX
PF   04-OCT-2000; 2000WO-US27464.
XX
PR   04-OCT-1999;   99US-0157455.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Xu J, Skelky YAW, Reed SG, Cheever MA;
XX
DR   WPI; 2001-245062/25.
DR   P-PSDB; AAB74817.
XX
PT   Prostate specific protein and its encoding polynucleotide, useful for
PT   the treatment and diagnosis of prostate cancer -
XX
PS   Claim 50; Page 233; 276pp; English.
XX
CC   The present invention describes an isolated polypeptide (I) comprising
CC   at least an immunogenic portion of a prostate tumour antigen protein or
CC   its variant. (I) have cytostatic activity and can be used in vaccine
CC   production. (I), prostate tumour antigen polynucleotides, an antigen
CC   presenting cell (APC e.g. a dendritic cell) that expresses (I), and a
CC   pharmaceutical composition containing (I) are useful for inhibiting the
CC   development of cancer in a patient. Antibodies specific for prostate
CC   specific proteins and oligonucleotides that hybridise to a
CC   polynucleotide that encodes a prostate specific protein are useful
CC   for detecting the presence or absence of a cancer or monitoring the
CC   progression the progression of a cancer, especially prostate cancer
CC   AAH02422 to AAH2872, AAB74798 to AAB74821 and AAB74830 are sequences
CC   used in the exemplification of the present invention.
XX
SQ   Sequence 2040 BP; 716 A; 392 C; 500 G; 432 T; 0 other;

  Query Match             97.7%;  Score 1128;  DB 22;  Length 2040;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1128;  Conservative    0;  Mismatches    0;  Indels    0,  Gaps    0;

Qy       1 atggtgattgaaggttgattccatgccggctgcctcttctgtgaagaaagccatttggtctc 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 atggtgattgaaggttgattccatgccggctgcctcttctgtgaagaaagccatttggtctc 60

Qy      61 agcagcaagatgggcaagtggtgctgccgttgcttccccctgctgcagggagcagcggcaag 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 agcagcaagatgggcaagtggtgctgccgttgcttccccctgctgcagggagcagcggcaag 120

Qy     121 agcaacgtgggcacttctggagaccacgacgactctgctatgaagacactcaggaagcaag 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 agcaacgtgggcacttctggagaccacgacgactctgctatgaagacactcaggaagcaag 180

Qy     181 atgggcaagtggtgccgccactgcttccccctgctgcagggggagtggcaagagcaacgtg 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 atgggcaagtggtgccgccactgcttccccctgctgcagggggagtggcaagagcaacgtg 240

Qy     241 ggcgcttctggagaccacgacgactctgctatgaagacactcaggaacaagatgggcaag 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 ggcgcttctggagaccacgacgactctgctatgaagacactcaggaacaagatgggcaag 300

Qy     301 tggtgctgccactgcttccccctgctgcagggggagcggcaagagcaagatgggcgcttgg 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 tggtgctgccactgcttccccctgctgcagggggagcggcaagagcaagatgggcgcttgg 360

Qy     361 ggagactacgatgacagtgccttcatggagcccaggtaccacgtccgtgggagaaratctg 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 ggagactacgatgacagtgccttcatggagcccaggtaccacgtccgtgggagaaratctg 420

Qy     421 gacaagctccacagagctgcctggtgggtaaagtcccagaaaggatctcatcatcatg 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 gacaagctccacagagctgcctggtgggtaaagtcccagaaaggatctcatcatcatg 480

Qy     481 ctcagggacactgacgtgaacaagaaggacaagcaaaagaggactgctctacattctggcc 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 ctcagggacactgacgtgaacaagaaggacaagcaaaagaggactgctctacattctggcc 540

Qy     541 tctgccaatgggaattcagaagtagtaaaactcctgctgcagacagacgatgtcaattaat 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 tctgccaatgggaattcagaagtagtaaaactcctgctgcagacagacgatgtcaattaat 600

Qy     601 gtccttgacaacaaaaagagagacagctctgataaaagccctacaatgccagagaaratgaa 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 gtccttgacaacaaaaagagagacagctctgataaaagccctacaatgccagagaaratgaa 660

Qy     661 tgtgcgttaatgttgctgtgaaracatgcactgatccaaatattccagatgagttatgaaat 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 tgtgcgttaatgttgctgtgaaracatgcactgatccaaatattccagatgagttatgaaat 720

Qy     721 accactctgcactacgctatctataatgaagataaattaatgccaaaagcactgtctta 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 accactctgcactacgctatctataatgaagataaattaatgccaaaagcactgtctta 780

Qy     781 tatggtgctgatatcgaatcaaaaaacaagcatggctcacaccactgttacttggtgta 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 tatggtgctgatatcgaatcaaaaaacaagcatggctcacaccactgttacttggtgta 840

Qy     841 catgagcaaaaacagcaagtcgtgaaatttttaatcaagaaaaaaagcgaatttaaatgca 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 catgagcaaaaacagcaagtcgtgaaatttttaatcaagaaaaaaagcgaatttaaatgca 900

Qy     901 ctggatagatatggaaggactgctctcatacttgctgtatgttgtggatcagcaagtata 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 ctggatagatatggaaggactgctctcatacttgctgtatgttgtggatcagcaagtata 960

Qy     961 gtcagcgcttctacttgagcaaaatattgatgtatcttctcaagatctatctggacagacg 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 gtcagcgcttctacttgagcaaaatattgatgtatcttctcaagatctatctggacagacg 1020

Qy    1021 gccagagagtatgctgtttctagtcatcatcatgtaatttgccagttacttttctgactac 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 gccagagagtatgctgtttctagtcatcatcatgtaatttgccagttacttttctgactac 1080

Qy    1081 aaagaaaaacagatgctaaaaatctcttctgaaaacagcaatccagaa 1128
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 aaagaaaaacagatgctaaaaatctcttctgaaaacagcaatccagaa 1128


RESULT  15
AAA06599
ID   AAA06599 standard; cDNA; 2000 BP.
XX
AC   AAA06599;
XX
DT   13 JUN 2000  (first entry)
XX
DE   Human immunogenic prostate tumour protein cDNA sequence SEQ ID NO:374.
XX
KW   Human; prostate cancer; diagnosis; tumour; gene therapy; detection;
KW   immunogenic; cytostatic; vaccine; ss.
XX
OS   Homo sapiens.
XX
PN   WO200004149-A2.
```

```
XX
PD   27-JAN-2000.
XX
PF   14-JUL-1999;   99WO-US15838.
XX
PR   14-JUL-1998;   98US-0115453.
PR   14-JUL-1998;   98US-0116134.
PR   23-SEP-1998;   98US-0159812.
PR   23-SEP-1998;   98US-0159822.
PR   15-JAN-1999;   99US-0232149.
PR   15-JAN-1999;   99US-0232880.
PR   09-APR-1999;   99US-0288946.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Dillon DC,  Harlocker SL,  Yuqiu J,  Xu J,  Mitcham JL;
XX
DR   WPI; 2000-171268/15.
XX
PT   New polypeptide useful for treating and diagnosing prostate cancer
PT   comprises an immunogenic portion of prostate tumor protein -
XX
PS   Claim 50; Page 222-223; 264pp; English.
XX
CC   The present invention describes isolated polypeptides, comprising an
CC   immunogenic portion of a prostate tumour protein (PTP). The polypeptides
CC   and polynucleotides encoding them have cytostatic activity and can be
CC   used in vaccines and in gene therapy. The polypeptides and
CC   polynucleotides encoding them, antigen presenting cells which express
CC   the polypeptides, antibodies against the polypeptides and vaccines
CC   comprising them can be used for inhibiting the development of prostate
CC   cancer in a patient. The polypeptides can be used to generate antibodies
CC   or anti-idiotypic antibodies for passive immuno therapy. A portion of
CC   the polynucleotides encoding the polypeptides can be used as a probe or
CC   to modulate the expression of the polypeptides. AAA06241 to AAA06691 and
CC   AAY82000 to AAY82020 represent sequences used in the exemplification of
CC   the present invention.
XX
SQ   Sequence 2000 BP; 608 A; 487 C; 480 G; 426 T; 0 other;

  Query Match             93.2%;  Score 1077;  DB 21;  Length 2000;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1127;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 atggtggttgaggttgattccatgccggctgcctcttctgtgaagaagccatttggtctc 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 atggtggttgaggttgattccatgccggctgcctcttctgtgaagaagccatttggtctc 60

Qy         61 gggagcaagatgggcaagtggtgctgccgttgcttcccctgctgcaggagaagcggcaag 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 gggagcaagatgggcaagtggtgctgccgttgcttcccctgctgcaggagaagcggcaag 120

Qy        121 agcaacagtgggcacttctggagaccacgacgactctgctatgaagacactcaggagcaag 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 agcaacagtgggcacttctggagaccacgacgactctgctatgaagacactcaggagcaag 180

Qy        181 atgggcaagtggtgccgccactgcttcccctgctgcaggggagtggcaagagcaacgtg 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 atgggcaagtggtgccgccactgcttcccctgctgcaggggagtggcaagagcaacgtg 240

Qy        241 gggcgcttctggagaccacgacgactctgctatgaagacactcaggaacaagatgggcaag 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 gggcgcttctggagaccacgacgactctgctatgaagacactcaggaacaagatgggcaag 300

Qy        301 tggtgctgccgctgcttcccctgctgcagggggagcggcaagagcaaggtgggcgcttgg 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 tggtgctgccgctgcttcccctgctgcagggggagcggcaagagcaaggtgggcgcttgg 360

Qy        361 gggactacgatgacagtgccttcatggagcccaggtaccacgtccgtggagaagatctg 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 gggactacgatgacagtgccttcatggagcccaggtaccacgtccgtggagaagatctg 420

Qy        421 gacaagctccacagagctgcctggtgggtaaagtcccagaaaaggattcatcgtcatg 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 gacaagctccacagagctgcctggtgggtaaagtcccagaaaaggattcatcgtcatg 480

Qy        481 ctcagggacactgacgtgaacaagaaggacaagcaaaagaggactgcttctcatctggcc 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ctcagggacactgacgtgaacaagaaggacaagcaaaagaggactgcttctcatctggcc 540

Qy        541 tctgccaatgggaattcagaagtagtaaaactcctgctggacagacgatgtcaacttaat 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 tctgccaatgggaattcagaagtagtaaaactcctgctggacagacgatgtcaacttaat 600

Qy        601 gtccttgacaacaaaaagaggacagctctgataaaggccgtacaatgccaggaagatgaa 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 gtccttgacaacaaaaagaggacagctctgataaaggccgtacaatgccaggaagatgaa 660

Qy        661 tgtgcgttaatgttgctggaacatggcactgatccaaatattccagatgagtatggaaat 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 tgtgcgttaatgttgctggaacatggcactgatccaaatattccagatgagtatggaaat 720

Qy        721 accactctgcactacgctatctataatgaagataaattaatggccaaagcactgctctta 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 accactctgcactacgctatctataatgaagataaattaatggccaaagcactgctctta 780

Qy        781 tatggtgctgatatcgaatcaaaaaacaagcatggcctcacaccactgttacttggtgta 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 tatggtgctgatatcgaatcaaaaaacaagcatggcctcacaccactgttacttggtgta 840

Qy        841 catgagcaaaaacagcaagtcgtgaaatttttaatcaagaaaaaagcgaatttaaatgca 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 catgagcaaaaacagcaagtcgtgaaatttttaatcaagaaaaaagcgaatttaaatgca 900

Qy        901 ctggatagatatggaaggactgctctcatacttgctgtatgttgtggatcagcaagtata 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 ctggatagatatggaaggactgctctcatacttgctgtatgttgtggatcagcaagtata 960

Qy        961 gtcagccttctacttgagcaaaatattgatgtatcttctcaagatctatctggacagacg 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 gtcagccttctacttgagcaaaatattgatgtatcttctcaagatctatctggacagacg 1020

Qy       1021 gccagagagtatgctgtttctagtcatcatcatgtaatttgccagttactttctgactac 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 gccagagagtatgctgtttctagtcatcatcatgtaatttgccagttactttctgactac 1080

Qy       1081 aaagaaaaacagatgctaaaaatctcttctgaaaacagcaatccaaga 1128
              |||||||||||||||||||||||||||||||||||||||||||||||:
Db       1081 aaagaaaaacagatgctaaaaatctcttctgaaaacagcaatccaaga 1128


Search completed: March 9, 2002, 05:01:38
Job time: 6208 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: March 9, 2002, 03:06:47 ; Search time 78.91 Seconds
(without alignments)
3314.942 Million cell updates/sec

Title: US-09-699-295-301
Perfect score: 1155
Sequence: 1 atggatggttgaggttgattc..........caagaaccagaaataaataa 1155

Scoring table: IDENTITY_NUC
Gapop 60.0 , Gapext 60.0

Searched: 351203 seqs, 113238999 residues

Word size : 0

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	120	10.4	1851	4	US-08-991-789A-291	Sequence 291, App
2	110	9.4	454	4	US-08-991-789A-211	Sequence 211, App
3	18	1.6	3877	2	US-08-599-895-1	Sequence 1, Appli
4	18	1.6	3877	3	US-09-211-290-1	Sequence 1, Appli
5	18	1.6	3877	3	US-09-322-676-1	Sequence 1, Appli
6	18	1.6	3877	4	US-09-456-036A-1	Sequence 1, Appli
7	17	1.5	654	3	US-08-564-496C-18	Sequence 18, Appl
8	17	1.5	656	5	PCT-US95-15428-18	Sequence 18, Appl
9	17	1.5	744	1	US-08-650-129-1	Sequence 1, Appli
10	17	1.5	744	3	US-08-984-417-1	Sequence 1, Appli
11	17	1.5	821	1	US-08-650-129-2	Sequence 2, Appli
12	17	1.5	821	3	US-08-984-417-2	Sequence 2, Appli
13	17	1.5	866	1	US-08-650-129-3	Sequence 3, Appli
14	17	1.5	866	3	US-08-984-417-3	Sequence 3, Appli
15	17	1.5	2429	4	US-09-386-493-3	Sequence 3, Appli
16	17	1.5	2439	4	US-09-233-989-1	Sequence 1, Appli
17	17	1.5	2443	1	US-08-452-262-1	Sequence 1, Appli
18	17	1.5	2443	1	US-08-734-550-1	Sequence 1, Appli
19	17	1.5	2443	5	PCT-US96-02528-1	Sequence 1, Appli
20	17	1.5	40050	4	US-09-245-941-5	Sequence 5, Appli
21	16	1.4	85	1	US-08-447-169A-171	Sequence 171, App
22	16	1.4	86	1	US-08-447-169A-169	Sequence 169, App
23	16	1.4	228	6	5273901-10	Patent No. 5273901
24	16	1.4	228	6	5482789-9	Patent No. 5482789
25	16	1.4	378	2	US-08-875-972-8	Sequence 8, Appli
26	16	1.4	410	3	US-08-569-028-14	Sequence 14, Appl
27	16	1.4	410	3	US-08-784-582-14	Sequence 14, Appl
28	16	1.4	410	4	US-08-785-271-14	Sequence 14, Appl
29	16	1.4	431	2	US-08-345-321-7	Sequence 7, Appli
30	16	1.4	484	4	US-08-975-762-34	Sequence 34, Appl
31	16	1.4	484	4	US-08-821-324-34	Sequence 34, Appl
32	16	1.4	484	4	US-09-295-028-34	Sequence 34, Appl
33	16	1.4	484	4	US-09-306-582-34	Sequence 34, Appl
34	16	1.4	548	4	US-08-642-274B-21	Sequence 21, Appl
35	16	1.4	548	4	US-08-952-014C-21	Sequence 21, Appl
36	16	1.4	628	1	US-08-686-878A-1	Sequence 1, Appli
37	16	1.4	687	3	US-09-048-052-5	Sequence 5, Appli
38	16	1.4	687	4	US-09-160-246-5	Sequence 5, Appli
39	16	1.4	699	2	US-08-735-609-11	Sequence 11, Appl
40	16	1.4	699	2	US-08-735-609-11	Sequence 11, Appl
41	16	1.4	699	3	US-09-315-372-11	Sequence 11, Appl
42	16	1.4	699	3	US-09-244-752-11	Sequence 11, Appl
43	16	1.4	699	3	US-09-245-497-11	Sequence 11, Appl
44	16	1.4	848	1	US-08-686-878A-44	Sequence 44, Appl
45	16	1.4	985	4	US-09-160-246-9	Sequence 9, Appli

ALIGNMENTS

RESULT 1
US-08-991-789A-291/c
; Sequence 291, Application US/08991789A
; Patent No. 6225054
; GENERAL INFORMATION:
; APPLICANT: Frudakis, Tony N.
; Smith, John M.
; Reed, Steven G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TREATMENT AND DIAGNOSIS OF BREAST CANCER
; NUMBER OF SEQUENCES: 292
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed IP Law Group
; STREET: 701 Fifth Avenue, Suite 6300
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/991,789A
; FILING DATE: 11-Dec-1997
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Potter, Jane E. R.
; REGISTRATION NUMBER: 33,332
; REFERENCE/DOCKET NUMBER: 210121.41003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 291:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1851 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 291:
US-08-991-789A-291

Query Match 10.4%; Score 120; DB 4; Length 1851;
Best Local Similarity 99.1%; Pred. No. 1.6e-51;
Matches 220; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 507 ggacaagcaaaagaggactgctctacatctggcctctgccaatggagattcaaaatgatt 5..
|||||||||:|||||||||||||||||||||||||||||||||||||||||:|||||||||

```
Db    791 GGACAAGCAAAAGAGGACTGCTCTACATCTGGCCTCTGCCAATGGGAATTCAGAAGTAGT 732

Qy    567 aaaactctgctggacagacaatgtcaacttaatgtcttgacaacaaaaacaggacagc 626
          ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    731 AAAACTCGTGCTGGACAGACGATGTCAACTTAATGTCCTTGACAACAAAAACAGGACAGC 672

Qy    627 tctgatccaaaatattccagatgagtatggaaatacccactct 668
          ||||| |||||||||||||||||||||||||||||||||||||
Db    671 TCTGACAAAGGCCGTACAATGCCAGGAAGATGAATGTCGTTAATGTTGCTGGAACATGC 612

Qy    687 cactgatccaaaatattccagatgagtatggaaatacccactct 728
          |||||||||||||||||||||||||||||||||||||||||||
Db    611 CACTGATCCAAATATTCCAGATGAGTATGGAAATACCACTCT 570


RESULT   2
US-08-991-789A-211/c
; Sequence 211, Application US/08991789A
; Patent No. 6225054
;   GENERAL INFORMATION:
;        APPLICANT: Prudakis, Tony N.
;                   Smith, John M.
;                   Reed, Steven G.
;        TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
;                            TREATMENT AND DIAGNOSIS OF BREAST CANCER
;        NUMBER OF SEQUENCES: 292
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Seed IP Law Group
;             STREET: 701 Fifth Avenue, Suite 6300
;             CITY: Seattle
;             STATE: Washington
;             COUNTRY: USA
;             ZIP: 98104-7092
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/08/991,789A
;             FILING DATE: 11-Dec-1997
;             CLASSIFICATION: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Potter, Jane E. R.
;             REGISTRATION NUMBER: 33,332
;             REFERENCE/DOCKET NUMBER: 210121.419C3
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (206) 622-4900
;             TELEFAX: (206) 682-6031
;   INFORMATION FOR SEQ ID NO: 211:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 454 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 211:
US-08-991-789A-211


  Query Match             9.4%;  Score 109;  DB 4;  Length 454;
  Best Local Similarity   99.1%;  Pred. No. 5.7e 46;
  Matches  209;  Conservative    0;  Mismatches    2;  Indels    0,  Gaps    0,

Qy    507 qgacaagcaaaagaggactgctctacatctggcctctgccaatgggaattcagaagtagt 566
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    418 GGACAAGCAAAAGAGGACTGCTCTACATCTGGCCTCTGCCAATGGGAATTCAGAAGTAGT 359

Qy    567 aaaactctgctggacagacaatgtcaacttaatgtcttgacaacaaaaacaggacagc 626
          ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    358 AAAACTCGTGCTGGACAGACGATGTCAACTTAATGTCCTTGACAACAAAAACAGGACAGC 299

Qy    627 tctgatccaaaatattccagatgagtatggaaatacccactct 668
          ||||| |||||||||||||||||||||||||||||||||||||
Db    298 TCTGACAAAGGCCGTACAATGCCAGGAAGATGAATGTCGTTAATGTTGCTGGAACATGC 239

Qy    687 cactgatccaaaatattccagatgagtatggaa 717
          |||||||||||||||||||||||||||||||||
Db    238 CACTGATCCAAATATTCCAGATGAGTATGGAA 208


RESULT   3
US-08-599-895-1
; Sequence 1, Application US/08599895
; Patent No. 5891855
;  GENERAL INFORMATION:
;    APPLICANT:  Florkiewicz, Robert Z.
;    TITLE OF INVENTION:  INHIBITORS OF LEADERLESS PROTEIN EXPORT
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/599,895
;      FILING DATE:  31-JAN-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5891855tenburg Ph.D., Carol
;      REGISTRATION NUMBER:  39,817
;      REFERENCE/DOCKET NUMBER:  760100.416
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3877 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-599-895-1


  Query Match             1.6%;  Score 18;  DB 2;  Length 3877;
  Best Local Similarity   100.0%;  Pred. No. 12;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    863 tqaaattttaatcaaga 880
          ||||||||||||||||||
Db   3595 TGAAATTTTTAATCAAGA 3612


RESULT   4
US-09-211-290-1
; Sequence 1, Application US/09211290
; Patent No. 6071885
;  GENERAL INFORMATION:
;    APPLICANT:  Florkiewicz, Robert Z.
;    TITLE OF INVENTION:  INHIBITORS OF LEADERLESS PROTEIN EXPORT
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
```

```
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC DOS/MS DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/211,290
;       FILING DATE:  12 DEC 1998
;       CLASSIFICATION:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Maki, David J.
;       REGISTRATION NUMBER:  31,392
;       REFERENCE/DOCKET NUMBER:  200124.401D1
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (206) 622-4900
;       TELEFAX:  (206) 682-6031
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  3877 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
US-09-211-290-1


  Query Match                 1.6%;  Score 18;  DB 3;  Length 3877;
  Best Local Similarity   100.0%;  Pred. No. 12;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     863 tgaaattttaatcaaga 880
           ||||||||||||||||||
Db    3595 TGAAATTTTAATCAAGA 3612


RESULT   5
US-09-322-676-1
; Sequence 1, Application US/09322676
; Patent No. 6107283
;  GENERAL INFORMATION:
;    APPLICANT:  Florkiewicz, Robert Z.
;    TITLE OF INVENTION:  INHIBITORS OF LEADERLESS PROTEIN EXPORT
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/322,676
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 09/211,290
;      FILING DATE:  12 DEC 1998
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maki, David J.
;      REGISTRATION NUMBER:  31,392
;      REFERENCE/DOCKET NUMBER:  200124.401D1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3877 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-322-676-1


  Query Match                 1.6%;  Score 18;  DB 3;  Length 3877;
  Best Local Similarity   100.0%;  Pred. No. 12;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     863 tgaaattttaatcaaga 880
           ||||||||||||||||||
Db    3595 TGAAATTTTAATCAAGA 3612


RESULT   6
US-09-466-036A-1
; Sequence 1, Application US/09466036A
; Patent No. 6281197
;  GENERAL INFORMATION:
;       APPLICANT: Florkiewicz, Robert Z.
;       TITLE OF INVENTION: INHIBITORS OF LEADERLESS PROTEIN EXPORT
;       NUMBER OF SEQUENCES: 13
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: SEED and BERRY LLP
;            STREET: 6300 Columbia Center, 701 Fifth Avenue
;            CITY: Seattle
;            STATE: Washington
;            COUNTRY: USA
;            ZIP: 98104-7092
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC DOS/MS DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/466,036A
;            FILING DATE: 17-Dec-2001
;            CLASSIFICATION, <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US 09/211,290
;            FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Maki, David J.
;            REGISTRATION NUMBER: 31,392
;            REFERENCE/DOCKET NUMBER: 200124.401D1
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (206) 622-4900
;            TELEFAX: (206) 682-6031
;   INFORMATION FOR SEQ ID NO: 1:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 3877 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-466-036A-1


  Query Match                 1.6%;  Score 18;  DB 4;  Length 3877;
  Best Local Similarity   100.0%;  Pred. No. 12;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     863 tgaaattttaatcaaga 880
           ||||||||||||||||||
Db    3595 TGAAATTTTAATCAAGA 3612


RESULT   7
US-08-564-496C-18
; Sequence 18, Application US/08564496C
; Patent No. 6022687
;  GENERAL INFORMATION:
;    APPLICANT:  Lotarte, Michelle
;    APPLICANT:  Marchuk, Douglas A.
;    APPLICANT:  McAllister, Kimberly
;    TITLE OF INVENTION:  DIAGNOSIS OF AND THERAPY FOR
```

```
;    TITLE OF INVENTION:  HEREDITARY HAEMORRHAGIC TELANGIECTASIA
;    NUMBER OF SEQUENCES:  42
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/564,496C
;      FILING DATE:  29-NOV-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/346,129
;      FILING DATE:  29-NOV-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fraser, Ph.D., J.D., Janis K.
;      REGISTRATION NUMBER:  34,819
;      REFERENCE/DOCKET NUMBER:  06765/006001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617/542-5070
;      TELEFAX:  617/542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  18:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  654 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Genomic DNA
US-08-564-496C-18


  Query Match             1.5%;  Score 17;  DB 3;  Length 654;
  Best Local Similarity   100.0%;  Pred. No. 34;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     479 tgctcaqqqacactgac 495
           |||||||||||||||||
Db     108 TGCTCAGGGACACTGAC 124


RESULT   8
PCT-US95-15428-18
; Sequence 18, Application PC/TUS9515428
;  GENERAL INFORMATION:
;    APPLICANT:  Letarte, Michelle
;    APPLICANT:  Marchuk, Douglas A.
;    APPLICANT:  McAllister, Kimberly
;    TITLE OF INVENTION:  DIAGNOSIS OF AND THERAPY FOR
;    TITLE OF INVENTION:  HEREDITARY HAEMORRHAGIC TELANGIECTASIA
;    NUMBER OF SEQUENCES:  41
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  225 Franklin Street  Suite 3100
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30B
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/15428
;      FILING DATE:  29-NOV-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/346,129
;      FILING DATE:  29-NOV-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fraser, Janis K.
;      REGISTRATION NUMBER:  34,819
;      REFERENCE/DOCKET NUMBER:  06765/006001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617/542-5070
;      TELEFAX:  617/542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  18:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  656 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
PCT-US95-15428-18


  Query Match             1.5%;  Score 17;  DB 5;  Length 656;
  Best Local Similarity   100.0%;  Pred. No. 34;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     479 tgctcaqqqacactgac 495
           |||||||||||||||||
Db     111 TGCTCAGGGACACTGAC 127


RESULT   9
US-08-650-129-1/c
; Sequence 1, Application US/08650129
; Patent No. 5747322
;  GENERAL INFORMATION:
;    APPLICANT:  Craik, Charles S.
;    APPLICANT:  Tsu, Christopher A.
;    TITLE OF INVENTION:  Recombinant Crab Collagenase
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Majestic Parsons Siebert & Hsue
;      STREET:  Four Embarcadero Center, Suite 1450
;      CITY:  San Francisco
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94111-4121
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/650,129
;      FILING DATE:  09-MAY-1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Siebert, J. Suzanne
;      REGISTRATION NUMBER:  28,758
;      REFERENCE/DOCKET NUMBER:  2500.084US0
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 362-5556
;      TELEFAX:  (415) 362-5418
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  734 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-650-129-1


  Query Match             1.5%;  Score 17;  DB 1;  Length 734;
  Best Local Similarity   100.0%;  Pred. No. 35;
```

```
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     433 agagctgcctggtgggg 449
           |||||||||||||||||
Db      53 AGAGCTGCCTGGTGGGG 37


RESULT 10
US-08-984-417-1/c
; Sequence 1, Application US/08984417
; Patent No. 6017747
;  GENERAL INFORMATION:
;    APPLICANT:  Craik, Charles S.
;    APPLICANT:  Tsu, Christopher A.
;    TITLE OF INVENTION:  Recombinant Crab Collagenase
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Majestic Parsons Siebert & Hsue
;      STREET:  Four Embarcadero Center, Suite 1450
;      CITY:  San Francisco
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94111 4121
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/984,417
;      FILING DATE:  09 MAY 1996
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Siebert, J. Suzanne
;      REGISTRATION NUMBER:  28,758
;      REFERENCE/DOCKET NUMBER:  2500.084US0
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 362 5556
;      TELEFAX:  (415) 362 5418
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  734 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-984-417-1


  Query Match             1.5%;  Score 17;  DB 3;  Length 734;
  Best Local Similarity   100.0%;  Pred. No. 35;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     433 agagctgcctggtgggg 449
           |||||||||||||||||
Db      53 AGAGCTGCCTGGTGGGG 37


RESULT 11
US-08-650-129-2/c
; Sequence 2, Application US/08650129
; Patent No. 5747322
;  GENERAL INFORMATION:
;    APPLICANT:  Craik, Charles S.
;    APPLICANT:  Tsu, Christopher A.
;    TITLE OF INVENTION:  Recombinant Crab Collagenase
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Majestic Parsons Siebert & Hsue
;      STREET:  Four Embarcadero Center, Suite 1450
;      CITY:  San Francisco
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94111 4121
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/650,129
;      FILING DATE:  09-MAY-1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Siebert, J. Suzanne
;      REGISTRATION NUMBER:  28,758
;      REFERENCE/DOCKET NUMBER:  2500.084US0
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 362 5556
;      TELEFAX:  (415) 362 5418
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  821 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-650-129-2

  Query Match             1.5%;  Score 17;  DB 1;  Length 821;
  Best Local Similarity   100.0%;  Pred. No. 35;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     433 agagctgcctggtgggg 449
           |||||||||||||||||
Db     140 AGAGCTGCCTGGTGGGG 124


RESULT 12
US-08-984-417-2/c
; Sequence 2, Application US/08984417
; Patent No. 6017747
;  GENERAL INFORMATION:
;    APPLICANT:  Craik, Charles S.
;    APPLICANT:  Tsu, Christopher A.
;    TITLE OF INVENTION:  Recombinant Crab Collagenase
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Majestic Parsons Siebert & Hsue
;      STREET:  Four Embarcadero Center, Suite 1450
;      CITY:  San Francisco
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94111 4121
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/984,417
;      FILING DATE:  09 MAY 1996
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Siebert, J. Suzanne
;      REGISTRATION NUMBER:  28,758
;      REFERENCE/DOCKET NUMBER:  2500.084US0
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 362 5556
;      TELEFAX:  (415) 362 5418
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  821 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
```

```
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
US-08-984-417-2


  Query Match              1.5%;  Score 17;  DB 3;  Length 821;
  Best Local Similarity   100.0%;  Pred. No. 35;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     433 agagctgcctggtgggg 449
           |||||||||||||||||
Db     140 AGAGCTGCCTGGTGGGG 124


RESULT  13
US-08-650-129-3/c
; Sequence 3, Application US/08650129
; Patent No. 5747322
;  GENERAL INFORMATION:
;     APPLICANT:  Craik, Charles S.
;     APPLICANT:  Tsu, Christopher A.
;     TITLE OF INVENTION:  Recombinant Crab Collagenase
;     NUMBER OF SEQUENCES:  6
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Majestic Parsons Siebert & Hsue
;       STREET:  Four Embarcadero Center, Suite 1450
;       CITY:  San Francisco
;       STATE:  CA
;       COUNTRY:  USA
;       ZIP:  94111-4121
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/650,129
;       FILING DATE:  09-MAY-1996
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Siebert, J. Suzanne
;       REGISTRATION NUMBER:  28,758
;       REFERENCE/DOCKET NUMBER:  2500.084US0
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (415) 362 5556
;       TELEFAX:  (415) 362-5418
;   INFORMATION FOR SEQ ID NO:  3:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  866 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
US-08-650-129-3


  Query Match              1.5%;  Score 17;  DB 1;  Length 866;
  Best Local Similarity   100.0%;  Pred. No. 35;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     433 agagctgcctggtgggg 449
           |||||||||||||||||
Db     185 AGAGCTGCCTGGTGGGG 169


RESULT  14
US-08-984-417-3/c
; Sequence 3, Application US/08984417
; Patent No. 6017747
;  GENERAL INFORMATION:
;     APPLICANT:  Craik, Charles S.
;     APPLICANT:  Tsu, Christopher A.
;     TITLE OF INVENTION:  Recombinant Crab Collagenase
;     NUMBER OF SEQUENCES:  6
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Majestic Parsons Siebert & Hsue
;       STREET:  Four Embarcadero Center, Suite 1450
;       CITY:  San Francisco
;       STATE:  CA
;       COUNTRY:  USA
;       ZIP:  94111 4121
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-Dos/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/984,417
;       FILING DATE:  09-MAY-1996
;       CLASSIFICATION:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Siebert, J. Suzanne
;       REGISTRATION NUMBER:  28,758
;       REFERENCE/DOCKET NUMBER:  2500.084US0
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (415) 362-5556
;       TELEFAX:  (415) 362-5418
;   INFORMATION FOR SEQ ID NO:  3:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  866 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
US-08-984-417-3


  Query Match              1.5%;  Score 17;  DB 3;  Length 866;
  Best Local Similarity   100.0%;  Pred. No. 35;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     433 agagctgcctggtgggg 449
           |||||||||||||||||
Db     185 AGAGCTGCCTGGTGGGG 169


RESULT  15
US-09-386-493-3/c
; Sequence 3, Application US/09386493
; Patent No. 6262247
; GENERAL INFORMATION:
;  APPLICANT: Kaser, Matthew R.
;  APPLICANT: Azimzai, Yalda
;  APPLICANT: Yue, Henry
;  TITLE OF INVENTION: POLYCYCLIC AROMATIC HYDROCARBON INDUCED MOLECULES
;  FILE REFERENCE: PB 0011 US
;  CURRENT APPLICATION NUMBER: US/09/386,493
;  CURRENT FILING DATE: 1999-08-30
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: PERL Program
; SEQ ID NO 3
;   LENGTH: 2429
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: 2404, 2413, 2423, 2425, 2426
;   OTHER INFORMATION: a or g or c or t, unknown, or other
;   FEATURE:
;   NAME/KEY:
;   OTHER INFORMATION: 253053
;  PUBLICATION INFORMATION:
US-09-386-493-3


  Query Match              1.5%;  Score 17;  DB 4;  Length 2429;
```

Best Local Similarity 100.0%; Pred. No. 47;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    1101 ctcttctaaaacagca 1120
           ||||||||||||||||
Db     570 CTCTTCTGAAAACAGCA 554
```

Search completed: March 9, 2002, 04:13:14
Job time : 9087 sec

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: March 9, 2002, 09:31:21 ; Search time 1719.35 Seconds
(without alignments)
11082.236 Million cell updates/sec

Title: US-09-699-295-301
Perfect score: 1155
Sequence: 1 atggtggttgaagttgattc..........caagaaccagaaataaataa 1155

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_om:*
20: em_or:*
21: em_ov:*
22: em_pat:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_sts:*
27: em_sy:*
28: em_un:*
29: em_vi:*
30: em_htgo_hum:*
31: em_htgo_inv:*
32: em_htgo_rod:*
33: em_htg_hum:*
34: em_htg_inv:*
35: em_htg_rod:*
36: em_htg_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	Query Match	Length	DB	ID	Description
	1	1155	100.0	1155	6	AX106592	AX106592 Sequence
	2	1155	100.0	1155	6	AX140883	AX140883 Sequence
	3	1131.6	98.0	2000	6	AX106593	AX106593 Sequence
	4	1131.6	98.0	2000	6	AX140884	AX140884 Sequence
	5	1131.6	98.0	2040	6	AX106594	AX106594 Sequence
	6	1131.6	98.0	2040	6	AX140885	AX140885 Sequence
	7	1088.6	94.3	1853	6	AX106588	AX106588 Sequence
	8	1088.6	94.3	1853	6	AX140879	AX140879 Sequence
	9	841	72.8	1852	6	AX141040	AX141040 Sequence
c	10	829	71.8	1851	6	AR148112	AR148112 Sequence
c	11	829	71.8	1851	6	AX106585	AX106585 Sequence
c	12	829	71.8	1851	6	AX140876	AX140876 Sequence
	13	829	71.8	2184	6	AX106589	AX106589 Sequence
	14	829	71.8	2184	6	AX140880	AX140880 Sequence
	15	804.8	69.7	1512	6	AX106587	AX106587 Sequence
	16	804.8	69.7	1512	6	AX140878	AX140878 Sequence
	17	768	66.5	879	6	AX141041	AX141041 Sequence
	18	571.4	49.5	1855	6	AX106590	AX106590 Sequence
	19	571.4	49.5	1855	6	AX140881	AX140881 Sequence
	20	520	45.0	1059	6	AX106591	AX106591 Sequence
	21	520	45.0	1059	6	AX140882	AX140882 Sequence
c	22	518.8	44.9	199668	2	AC018804	AC018804 Homo sapi
c	23	480.4	41.6	154077	9	CNS07EFD	AL512624 Human chr
	24	478.8	41.5	115727	2	AC013269	AC013269 Homo sapi
	25	459.6	39.8	172550	2	AC060814	AC060814 Homo sapi
	26	456.4	39.5	153077	9	AC020679	AC020679 Homo sapi
	27	454.4	39.3	80955	2	AC026984	AC026984 Homo sapi
	28	451.6	39.1	104785	9	AP001465	AP001465 Homo sapi
	29	451.6	39.1	340000	9	HS21C003	AL163203 Homo sapi
c	30	417.8	36.2	454	6	AR148032	AR148032 Sequence
	31	377.8	32.7	64976	2	AC026256	AC026256 Homo sapi
	32	354.4	30.7	1805	9	AK001137	AK001137 Homo sapi
	33	354.4	30.7	5360	9	AB028997	AB028997 Homo sapi
	34	322	27.9	83276	2	AC026631	AC026631 Homo sapi
	35	322	27.9	184109	2	AC023191	AC023191 Homo sapi
c	36	322	27.9	217377	2	AC037471	AC037471 Homo sapi
	37	320.4	27.7	167293	2	AC022629	AC022629 Homo sapi
c	38	320.4	27.7	177317	2	AC036207	AC036207 Homo sapi
	39	299.4	25.9	3527	9	HSM801761	AL136799 Homo sapi
	40	278.4	24.1	4458	9	AF269087	AF269087 Homo sapi
	41	274.6	23.8	3673	9	AF269088	AF269088 Homo sapi
c	42	215.8	18.7	173957	2	AC091671	AC091671 Papio cyn
	43	206.8	17.9	1388	10	AF294328	AF294328 Mus muscu
	44	196	17.0	1067	9	D78334	D78334 Human mRNA
c	45	196	17.0	3210	9	AK024286	AK024286 Homo sapi

ALIGNMENTS

RESULT 1
AX106592
LOCUS AX106592 1155 bp DNA PAT 30-APR-2001
DEFINITION Sequence 373 from Patent WO0125272.
ACCESSION AX106592
VERSION AX106592.1 GI:13922263
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1155)
AUTHORS Xu,J., Skeiky,Y.A., Reed,S.G. and Cheever,M.A.
TITLE Compositions and methods for therapy and diagnosis of prostate cancer
JOURNAL Patent: WO 0125272-A 373 12 APR 2001;
CORIXA CORPORATION (US)
FEATURES Location/Qualifiers
source 1..1155
/organism "Homo sapiens"

```
                     /db_xref="taxon:9606"
BASE COUNT      346 a    253 c    297 g    259 t
ORIGIN

  Query Match             100.0%;  Score 1155;  DB 6;  Length 1155;
  Best Local Similarity   100.0%;  Pred. No. 4.5e-283;
  Matches 1155;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 atggtggttgaggttgattccatgccggctgcctcttctgtgaagaagccatttggtctc 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGTGGTTGAGGTTGATTCCATGCCGGCTGCCTCTTCTGTGAAGAAGCCATTTGGTCTC 60

Qy         61 aggagcaagatgggcaagtggtgctgccgttgcttcccctgctgcagggagagcggcaag 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGGAGCAAGATGGGCAAGTGGTGCTGCCGTTGCTTCCCCTGCTGCAGGGAGAGCGGCAAG 120

Qy        121 agcaacgtgggcacttctggagaccacgacgactctgctatgaagacactcaggagcaag 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGCAACGTGGGCACTTCTGGAGACCACGACGACTCTGCTATGAAGACACTCAGGAGCAAG 180

Qy        181 atgggcaagtggtgccgccactgcttcccctgctgcagggggagtggcaagagcaacgtg 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATGGGCAAGTGGTGCCGCCACTGCTTCCCCTGCTGCAGGGGGAGTGGCAAGAGCAACGTG 240

Qy        241 ggcgcttctggagaccacgacgactctgctatgaagacactcaggaacaagatgggcaag 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GGCGCTTCTGGAGACCACGACGACTCTGCTATGAAGACACTCAGGAACAAGATGGGCAAG 300

Qy        301 tggtgctgccactgcttcccctgctgcagggggagcggcaagagcaaggtgggcgcttgg 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TGGTGCTGCCACTGCTTCCCCTGCTGCAGGGGGAGCGGCAAGAGCAAGGTGGGCGCTTGG 360

Qy        361 ggagactacgatgacagtgccttcatggagcccaggtaccacgtccgtggagaagatctg 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GGAGACTACGATGACAGTGCCTTCATGGAGCCCAGGTACCACGTCCGTGGAGAAGATCTG 420

Qy        421 gacaagctccacagagctgcctggtggggtaaagtccccagaaaggatctcatcgtcatg 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GACAAGCTCCACAGAGCTGCCTGGTGGGGTAAAGTCCCCAGAAAGGATCTCATCGTCATG 480

Qy        481 ctcagggacactgacgtgaacaagaaggacaagcaaaagaggactgctctacatctggcc 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTCAGGGACACTGACGTGAACAAGAAGGACAAGCAAAAGAGGACTGCTCTACATCTGGCC 540

Qy        541 tctgccaatgggaattcagaagtagtaaaactcctgctggacagacgatgtcaacttaat 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TCTGCCAATGGGAATTCAGAAGTAGTAAAACTCCTGCTGGACAGACGATGTCAACTTAAT 600

Qy        601 gtccttgacaacaaaaagaggacagctctgataaaggccgtacaatgccaggaagatgaa 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GTCCTTGACAACAAAAAGAGGACAGCTCTGATAAAGGCCGTACAATGCCAGGAAGATGAA 660

Qy        661 tgtgcgttaatgttgctggaacatggcactgatccaaatattccagatgagtatggaaat 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGTGCGTTAATGTTGCTGGAACATGGCACTGATCCAAATATTCCAGATGAGTATGGAAAT 720

Qy        721 accactctgcactacgctatctataatgaagataaattaatggccaaagcactgctctta 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ACCACTCTGCACTACGCTATCTATAATGAAGATAAATTAATGGCCAAAGCACTGCTCTTA 780

Qy        781 tatggtgctgatatcgaatcaaaaaacaagcatggcctcacaccactgttacttggtgta 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TATGGTGCTGATATCGAATCAAAAAACAAGCATGGCCTCACACCACTGTTACTTGGTGTA 840

Qy        841 catgagcaaaaacagcaagtcgtgaaatttttaatcaagaaaaaagcgaatttaaatgca 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CATGAGCAAAAACAGCAAGTCGTGAAATTTTTAATCAAGAAAAAAGCGAATTTAAATGCA 900

Qy        901 ctggatagatatggaaggactgctctcatacttgctgtatgttgtggatcagcaagtata 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CTGGATAGATATGGAAGGACTGCTCTCATACTTGCTGTATGTTGTGGATCAGCAAGTATA 960

Qy        961 gtcagccttctacttgagcaaaatattgatgtatcttctcaagatctatctggacagacg 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GTCAGCCTTCTACTTGAGCAAAATATTGATGTATCTTCTCAAGATCTATCTGGACAGACG 1020

Qy       1021 gccagagagtatgctgtttctagtcatcatcatgtaatttgccagttactttctgactac 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GCCAGAGAGTATGCTGTTTCTAGTCATCATCATGTAATTTGCCAGTTACTTTCTGACTAC 1080

Qy       1081 aaagaaaaacagatgctaaaaatctcttctgaaaacagcaatccagaaaatgtctcaaga 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AAAGAAAAACAGATGCTAAAAATCTCTTCTGAAAACAGCAATCCAGAAAATGTCTCAAGA 1140

Qy       1141 accagaaataaataa 1155
              |||||||||||||||
Db       1141 ACCAGAAATAAATAA 1155

RESULT 2
AX140883
LOCUS       AX140883       1155 bp    DNA             PAT       31-MAY-2001
DEFINITION  Sequence 373 from Patent WO0134802.
ACCESSION   AX140883
VERSION     AX140883.1  GI:14280986
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1155)
  AUTHORS   Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,
            Reed,S.G., Kalos,M.D., Retter,M.W., Stolk,J.A., Day,C.H.,
            Skeiky,Y.A. and Wang,A.
  TITLE     Compositions and methods for the therapy and diagnosis of prostate
            cancer
  JOURNAL   Patent: WO 0134802 A 373 17 MAY 2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..1155
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      346 a    253 c    297 g    259 t
ORIGIN

  Query Match             100.0%;  Score 1155;  DB 6;  Length 1155;
  Best Local Similarity   100.0%;  Pred. No. 4.5e-283;
  Matches 1155;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 atggtggttgaggttgattccatgccggctgcctcttctgtgaagaagccatttggtctc 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGTGGTTGAGGTTGATTCCATGCCGGCTGCCTCTTCTGTGAAGAAGCCATTTGGTCTC 60

Qy         61 aggagcaagatgggcaagtggtgctgccgttgcttcccctgctgcagggagagcggcaag 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGGAGCAAGATGGGCAAGTGGTGCTGCCGTTGCTTCCCCTGCTGCAGGGAGAGCGGCAAG 120

Qy        121 agcaacgtgggcacttctggagaccacgacgactctgctatgaagacactcaggagcaag 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGCAACGTGGGCACTTCTGGAGACCACGACGACTCTGCTATGAAGACACTCAGGAGCAAG 180

Qy        181 atgggcaagtggtgccgccactgcttcccctgctgcagggggagtggcaagagcaacgtg 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATGGGCAAGTGGTGCCGCCACTGCTTCCCCTGCTGCAGGGGGAGTGGCAAGAGCAACGTG 240

Qy        241 ggcgcttctggagaccacgacgactctgctatgaagacactcaggaacaagatgggcaag 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GGCGCTTCTGGAGACCACGACGACTCTGCTATGAAGACACTCAGGAACAAGATGGGCAAG 300

Qy        301 tggtgctgccactgcttcccctgctgcagggggagcggcaagagcaaggtgggcgcttgg 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TGGTGCTGCCACTGCTTCCCCTGCTGCAGGGGGAGCGGCAAGAGCAAGGTGGGCGCTTGG 360

Qy        361 ggagactacgatgacagtgccttcatggagcccaggtaccacgtccgtggagaagatctg 420
```

```
Db     361 GGAGACTACGATGACAGTGCCTTCATGGAGCCAGGTACCACGTCCGTGGAGAAGATCTG 420

Qy     421 gacaagctccacagagctgcctggtggggtaaagtccccagaaaggatctcatcgtcatg 480
Db     421 GACAAGCTCCACAGAGCTGCCTGGTGGGGTAAAGTCCCCAGAAAGGATCTCATCGTCATG 480

Qy     481 ctcagggacactgacgtgaacaagaaggacaagcaaaagaggactgctctacatctggcc 540
Db     481 CTCAGGGACACTGACGTGAACAAGAAGGACAAGCAAAAGAGGACTGCTCTACATCTGGCC 540

Qy     541 tctgccaatgggaattcagaagtagtaaaactcctgctggacagacgatgtcaacttaat 600
Db     541 TCTGCCAATGGGAATTCAGAAGTAGTAAAACTCCTGCTGGACAGACGATGTCAACTTAAT 600

Qy     601 gtccttgacaacaaaaagaggacagctctgataaaggccgtacaatgccaggaagatgaa 660
Db     601 GTCCTTGACAACAAAAAGAGGACAGCTCTGATAAAGGCCGTACAATGCCAGGAAGATGAA 660

Qy     661 tgtgcgttaatgttgctggaacatggcactgatccaaatattccagatgagtatggaaat 720
Db     661 TGTGCGTTAATGTTGCTGGAACATGGCACTGATCCAAATATTCCAGATGAGTATGGAAAT 720

Qy     721 accactctgcactacgctatctataatgaagataaattaatggccaaagcactgctctta 780
Db     721 ACCACTCTGCACTACGCTATCTATAATGAAGATAAATTAATGGCCAAAGCACTGCTCTTA 780

Qy     781 tatggtgctgatatcgaatcaaaaaacaagcatggcctcacaccactgttacttggtgta 840
Db     781 TATGGTGCTGATATCGAATCAAAAAACAAGCATGGCCTCACACCACTGTTACTTGGTGTA 840

Qy     841 catgagcaaaaacagcaagtcgtgaaatttttaatcaagaaaaaagcgaatttaaatgca 900
Db     841 CATGAGCAAAAACAGCAAGTCGTGAAATTTTTAATCAAGAAAAAAGCGAATTTAAATGCA 900

Qy     901 ctggatagatatggaaggactgctctcatacttgctgtatgttgtggatcagcaagtata 960
Db     901 CTGGATAGATATGGAAGGACTGCTCTCATACTTGCTGTATGTTGTGGATCAGCAAGTATA 960

Qy     961 gtcagccttctacttgagcaaaatattgatgtatcttctcaagatctatctggacagacg 1020
Db     961 GTCAGCCTTCTACTTGAGCAAAATATTGATGTATCTTCTCAAGATCTATCTGGACAGACG 1020

Qy    1021 gccagagagtatgctgtttctagtcatcatcatgtaatttgccagttactttctgactac 1080
Db    1021 GCCAGAGAGTATGCTGTTTCTAGTCATCATCATGTAATTTGCCAGTTACTTTCTGACTAC 1080

Qy    1081 aaagaaaaacagatgctaaaaatctcttctgaaaacagcaatccagaaaatgtctcaaga 1140
Db    1081 AAAGAAAAACAGATGCTAAAAATCTCTTCTGAAAACAGCAATCCAGAAAATGTCTCAAGA 1140

Qy    1141 accagaaataaataa 1155
Db    1141 ACCAGAAATAAATAA 1155
```

```
RESULT 4
AX106593
LOCUS       AX106593      2000 bp    DNA             PAT       30-APR-2001
DEFINITION  Sequence 374 from Patent WO0125272.
ACCESSION   AX106593
VERSION     AX106593.1  GI:13922264
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2000)
  AUTHORS   Xu,J., Skeiky,Y.A., Reed,S.G. and Cheever,M.A.
  TITLE     Compositions and methods for therapy and diagnosis of prostate
            cancer
  JOURNAL   Patent: WO 0125272-A 374 12-APR-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..2000
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      698 a    388 c    489 g    425 t
ORIGIN

  Query Match             98.0%;  Score 1131.6;  DB 6;  Length 2000;
  Best Local Similarity   99.6%;  Pred. No. 4.1e-277;
  Matches 1134;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 atggtggttgaggttgattccatgccggctgcctcttctgtgaagaagccatttggtctc 60
Db       1 ATGGTGGTTGAGGTTGATTCCATGCCGGCTGCCTCTTCTGTGAAGAAGCCATTTGGTCTC 60

Qy      61 aggagcaagatgggcaagtggtgctgccgttgcttcccctgctgcagggagagcggcaag 120
Db      61 AGGAGCAAGATGGGCAAGTGGTGCTGCCGTTGCTTCCCCTGCTGCAGGGAGAGCGGCAAG 120

Qy     121 agcaacgtgggcacttctggagaccacgacgactctgctatgaagacactcaggagcaag 180
Db     121 AGCAACGTGGGCACTTCTGGAGACCACGACGACTCTGCTATGAAGACACTCAGGAGCAAG 180

Qy     181 atgggcaagtggtgccgccactgcttcccctgctgcaggggcagtggcaagagcaacgtg 240
Db     181 ATGGGCAAGTGGTGCCGCCACTGCTTCCCCTGCTGCAGGGGCAGTGGCAAGAGCAACGTG 240

Qy     241 ggcgcttctggagaccacgacgactctgctatgaagacactcaggaacaagatgggcaag 300
Db     241 GGCGCTTCTGGAGACCACGACGACTCTGCTATGAAGACACTCAGGAACAAGATGGGCAAG 300

Qy     301 tggtgctgccactgcttcccctgctgcaggggagcggcaagagcaagtgggcgcttgg 360
Db     301 TGGTGCTGCCACTGCTTCCCCTGCTGCAGGGGAGCGGCAAGAGCAAGTGGGCGCTTGG 360

Qy     361 ggagactacgatgacagtgccttcatggagccaggtaccacgtccgtggagaagatctg 420
Db     361 GGAGACTACGATGACAGTGCCTTCATGGAGCCAGGTACCACGTCCGTGGAGAAGATCTG 420

Qy     421 gacaagctccacagagctgcctggtggggtaaagtccccagaaaggatctcatcgtcatg 480
Db     421 GACAAGCTCCACAGAGCTGCCTGGTGGGGTAAAGTCCCCAGAAAGGATCTCATCGTCATG 480

Qy     481 ctcagggacactgacgtgaacaagaaggacaagcaaaagaggactgctctacatctggcc 540
Db     481 CTCAGGGACACTGACGTGAACAAGAAGGACAAGCAAAAGAGGACTGCTCTACATCTGGCC 540

Qy     541 tctgccaatgggaattcagaagtagtaaaactcctgctggacagacgatgtcaacttaat 600
Db     541 TCTGCCAATGGGAATTCAGAAGTAGTAAAACTCCTGCTGGACAGACGATGTCAACTTAAT 600

Qy     601 gtccttgacaacaaaaagaggacagctctgataaaggccgtacaatgccaggaagatgaa 660
Db     601 GTCCTTGACAACAAAAAGAGGACAGCTCTGATAAAGGCCGTACAATGCCAGGAAGATGAA 660

Qy     661 tgtgcgttaatgttgctggaacatggcactgatccaaatattccagatgagtatggaaat 720
Db     661 TGTGCGTTAATGTTGCTGGAACATGGCACTGATCCAAATATTCCAGATGAGTATGGAAAT 720

Qy     721 accactctgcactacgctatctataatgaagataaattaatggccaaagcactgctctta 780
Db     721 ACCACTCTGCACTACGCTATCTATAATGAAGATAAATTAATGGCCAAAGCACTGCTCTTA 780

Qy     781 tatggtgctgatatcgaatcaaaaaacaagcatggcctcacaccactgttacttggtgta 840
Db     781 TATGGTGCTGATATCGAATCAAAAAACAAGCATGGCCTCACACCACTGTTACTTGGTGTA 840

Qy     841 catgagcaaaaacagcaagtcgtgaaatttttaatcaagaaaaaagcgaatttaaatgca 900
Db     841 CATGAGCAAAAACAGCAAGTCGTGAAATTTTTAATCAAGAAAAAAGCGAATTTAAATGCA 900

Qy     901 ctggatagatatggaaggactgctctcatacttgctgtatgttgtggatcagcaagtata 960
```

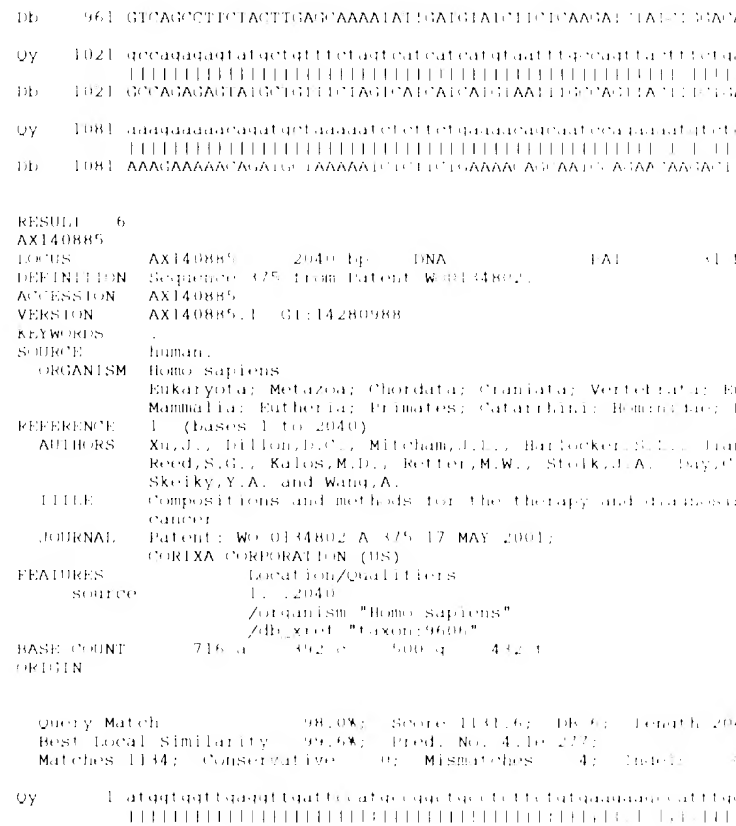
```
Db    901 CTGGATAGATATGGAAGGACTGCTCTCATACTTGCTGTATGTTGTGGATCAGCAAGTATA 960

Qy    961 gtcagccttctacttgaagcaaaatattgatgtatcttctcaagatctatctggacagacg 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 GTCAGCCTTCTACTTGAGCAAAATATTGATGTATCTTCTCAAGATCTATCTGGACAGACG 1020

Qy   1021 gccagaggagtatgctgtttctagtcatcatcatgtaatttgccagttacttttctgactac 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 GCCAGAGAGTATGCTGTTTCTAGTCATCATCATGTAATTTGCCAGTTACTTTTCTGACTAC 1080

Qy   1081 aaagaaaaacagatgctaaaaatctcttctgaaaacagcaatccagaaaatgtctcaa 1138
          |||||||||||||||||||||||||||||||||||||||||||||||| | | || ||
Db   1081 AAAGAAAAACAGATGCTAAAAATCTCTTCTGAAAACAGCAATCCAGAACAAGACTTAA 1138


RESULT   4
AX140884
LOCUS       AX140884     2000 bp    DNA             PAT       31-MAY-2001
DEFINITION  Sequence 374 from Patent WO0134802.
ACCESSION   AX140884
VERSION     AX140884.1  GI:14280987
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2000)
  AUTHORS   Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,
            Reed,S.G., Kalos,M.D., Retter,M.W., Stolk,J.A., Day,C.H.,
            Skeiky,Y.A. and Wang,A.
  TITLE     Compositions and methods for the therapy and diagnosis of prostate
            cancer
  JOURNAL   Patent: WO 0134802-A 374 17-MAY-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .2000
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      698 a    388 c    489 g    425 t
ORIGIN


  Query Match             98.0%;  Score 1131.6;  DB 6;  Length 2000;
  Best Local Similarity   99.6%;  Pred. No. 4.1e-277;
  Matches 1134;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 atggtggttgaggttgattccatgccggctgcctcttctgtgaagaagccatttggtctc 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ATGGTGGTTGAGGTTGATTCCATGCCGGCTGCCTCTTCTGTGAAGAAGCCATTTGGTCTC 60

Qy      61 aggaagcaagatgggcaaatggtgctgccgcttgcttccccctgctgcagggagagcggcaag 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 AGGAGCAAGATGGGCAAGTGGTGCTGCCGCTTGCTTCCCCTGCTGCAGGGAGAGCGGCAAG 120

Qy     121 agcaacgtgggcacttctggagaccacgacgactctgctatgaagacactcaggagcaag 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 AGCAACGTGGGCACTTCTGGAGACCACGACGACTCTGCTATGAAGACACTCAGGAGCAAG 180

Qy     181 atgggcaagtggtgccgccgcttgcttccccctgctgcagggggagtcggcaagagcaacgtg 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 ATGGGCAAGTGGTGCCGCCGCTTGCTTCCCCTGCTGCAGGGGGAGTCGGCAAGAGCAACGTG 240

Qy     241 ggcgcttctggaggaccacgacgactctgctatgaagacactcaggagcaagatgggcaag 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 GGCGCTTCTGGAGACCACGACGACTCTGCTATGAAGACACTCAGGAGCAAGATGGGCAAG 300

Qy     301 tggtgctgccgcttgcttccccctgctgcaggggggagcggcaaggagcaagtgggcgcttgg 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 TGGTGCTGCCGCTTGCTTCCCCTGCTGCAGGGGGAGCGGCAAGAGCAAGTGGGCGCTTGG 360

Qy     361 ggaggactacgatgacagtgccttcatggagcccaggtaccacgtccgtggagaagatctg 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 GGAGACTACGATGACAGTGCCTTCATGGAGCCCAGGTACCACGTCCGTGGAGAAGATCTG 420

Qy     421 gacaagctccacagagctgcctggtgggtaaagtcccctaaaggatctcatcgtcatg 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GACAAGCTCCACAGAGCTGCCTGGTGGGTAAAGTCCCCTAAAGGATCTCATCGTCATG 480

Qy     481 ctcagggacactgacgtgaacaagaaggacaagcaaaagaggactgctctacatctggcc 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 CTCAGGGACACTGACGTGAACAAGAAGGACAAGCAAAAGAGGACTGCTCTACATCTGGCC 540

Qy     541 tctgccaatgggaattcagaagtagtaaaactcctgctggacagacgatgtcaacttaat 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 TCTGCCAATGGGAATTCAGAAGTAGTAAAACTCCTGCTGGACAGACGATGTCAACTTAAT 600

Qy     601 gtccttgacaacaaaaaagaggacagctctgataaaggccgtacaatgccaggaagatgaa 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 GTCCTTGACAACAAAAAGAGGACAGCTCTGATAAAGGCCGTACAATGCCAGGAAGATGAA 660

Qy     661 tgtgcgttaatgttgctggaacatggcactgatccaaatattccagatgagtatgcaaat 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 TGTGCGTTAATGTTGCTGGAACATGGCACTGATCCAAATATTCCAGATGAGTATGCAAAT 720

Qy     721 accactctgcactacgctatctataatgaagataaattaatggccaaagcactgcttctta 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 ACCACTCTGCACTACGCTATCTATAATGAAGATAAATTAATGGCCAAAGCACTGCTTCTTA 780

Qy     781 tatggtgctgatatcgaatcaaaaaaacaagcatggcctcacaccactgttacttaatgta 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 TATGGTGCTGATATCGAATCAAAAAACAAGCATGGCCTCACACCACTGTTACTTAATGTA 840

Qy     841 catgagcaaaaacagcaagtcgtgaaatttttaatcaagaaaaaagcgaatttaaatgca 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 CATGAGCAAAAACAGCAAGTCGTGAAATTTTTAATCAAGAAAAAAGCGAATTTAAATGCA 900

Qy     901 ctggatagatatggaaggactgctctcatacttgctgtatgttgtggatcagcaagtata 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 CTGGATAGATATGGAAGGACTGCTCTCATACTTGCTGTATGTTGTGGATCAGCAAGTATA 960

Qy     961 gtcagccttctacttgaagcaaaatattgatgtatcttctcaagatctatctggacagacg 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 GTCAGCCTTCTACTTGAGCAAAATATTGATGTATCTTCTCAAGATCTATCTGGACAGACG 1020

Qy    1021 gccagaggagtatgctgtttctagtcatcatcatgtaatttgccagttacttttctgactac 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 GCCAGAGAGTATGCTGTTTCTAGTCATCATCATGTAATTTGCCAGTTACTTTTCTGACTAC 1080

Qy    1081 aaagaaaaacagatgctaaaaatctcttctgaaaacagcaatccagaaaatgtctcaa 1138
          |||||||||||||||||||||||||||||||||||||||||||||||| | | || ||
Db    1081 AAAGAAAAACAGATGCTAAAAATCTCTTCTGAAAACAGCAATCCAGAACAAGACTTAA 1138


RESULT   5
AX106594
LOCUS       AX106594     2040 bp    DNA             PAT       30-APR-2001
DEFINITION  Sequence 375 from Patent WO0125272.
ACCESSION   AX106594
VERSION     AX106594.1  GI:13922265
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2040)
  AUTHORS   Xu,J., Skeiky,Y.A., Reed,S.G. and Cheever,M.A.
  TITLE     Compositions and methods for therapy and diagnosis of prostate
            cancer
  JOURNAL   Patent: WO 0125272-A 375 12-APR-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .2040
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      716 a    392 c    500 g    432 t
```

ORIGIN

Query Match 98.0%; Score 1141.6; DB 6; Length 2040;
Best Local Similarity 99.6%; Pred. No. 4.1e-277;
Matches 1134; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

RESULT 6
AX140885
LOCUS AX140885 2040 bp DNA HAT 15-MAY-2001
DEFINITION Sequence 375 from Patent WO0134802.
ACCESSION AX140885
VERSION AX140885.1 GI:14280988
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2040)
AUTHORS Xu,J., Dillon,D.C., Mitcham,J.L., Burtscher,S.L., Tranel,Y.,
Reed,S.G., Kalos,M.D., Retter,M.W., Stolk,J.A., Day,C.B.,
Skolky,Y.A. and Wang,A.
TITLE Compositions and methods for the therapy and diagnosis of prostate
cancer
JOURNAL Patent: WO 0134802 A 375 17-MAY-2001;
CORIXA CORPORATION (US)
FEATURES Location/Qualifiers
source 1..2040
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 716 a 492 c 500 g 432 t
ORIGIN

Query Match 98.0%; Score 1141.6; DB 6; Length 2040;
Best Local Similarity 99.6%; Pred. No. 4.1e-277;
Matches 1134; Conservative 0; Mismatches 4; Indels 0; Gaps 0;



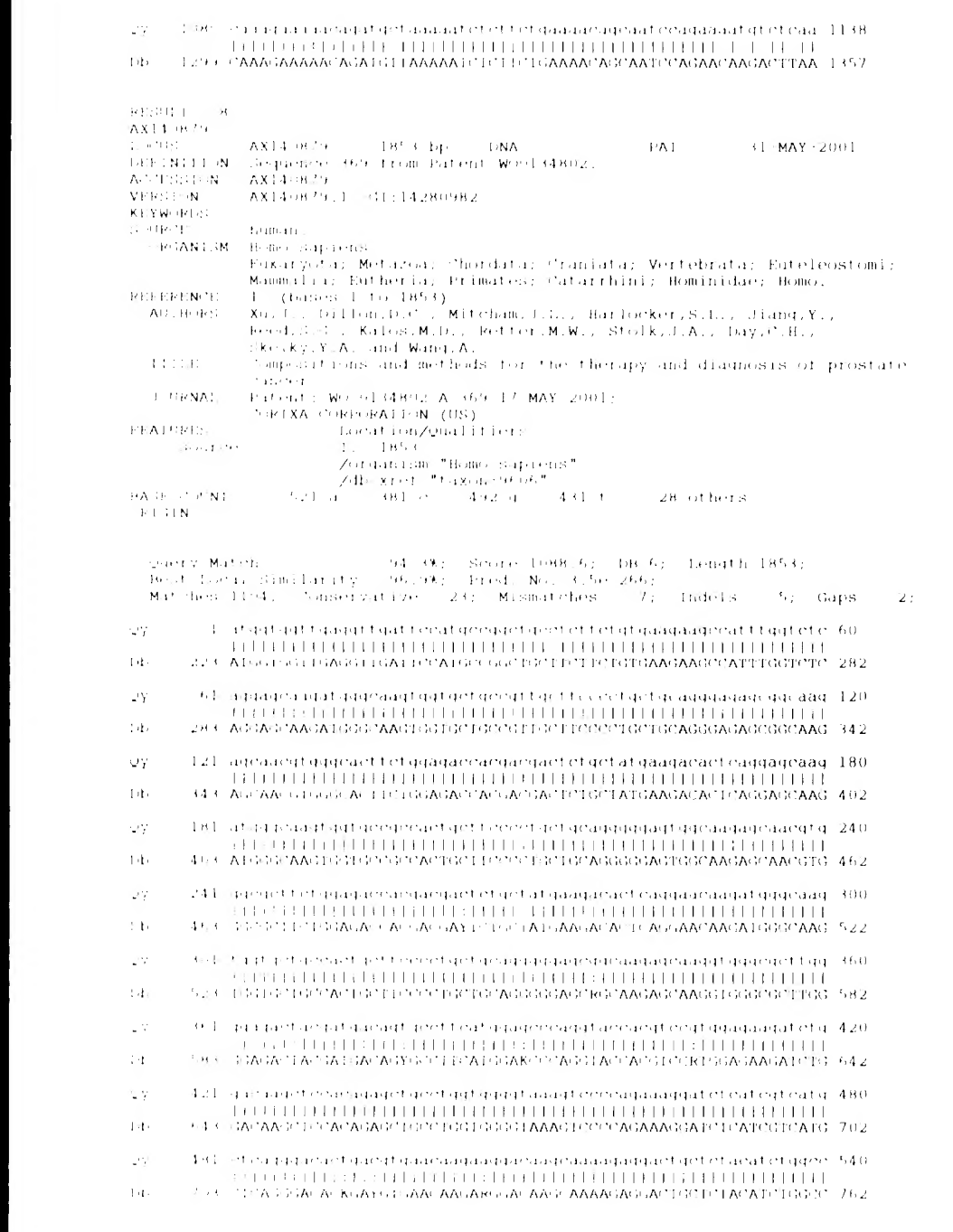
```
Qy     481 ctcagggacactgacgtgaacaagaaggacaagcaaaagaggactgctctacatctggcc 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 CTCAGGGACACTGACGTGAACAAGAAGGACAAGCAAAAGAGGACTGCTCTACATCTGGCC 540

Qy     541 tctgccaatgggaattcagaagtagtaaaactcctgctggacagacgatgtcaacttaat 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 TCTGCCAATGGGAATTCAGAAGTAGTAAAACTCCTGCTGGACAGACGATGTCAACTTAAT 600

Qy     601 gtccttgacaacaaaaagaggacagctctgataaaggccgtacaatgccaggaagatgaa 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 GTCCTTGACAACAAAAAGAGGACAGCTCTGATAAAGGCCGTACAATGCCAGGAAGATGAA 660

Qy     661 tgtgcgttaatgttgctggaacatggcactgatccaaatattccagatgagtatggaaat 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 TGTGCGTTAATGTTGCTGGAACATGGCACTGATCCAAATATTCCAGATGAGTATGGAAAT 720

Qy     721 accactctgcactacgctatctataatgaaqataaattaatggccaaaqcactgctctta 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 ACCACTCTGCACTACGCTATCTATAATGAAGATAAATTAATGGCCAAAGCACTGCTCTTA 780

Qy     781 tatggtgctgatatcgaatcaaaaaacaagcatggcctcacaccactgttacttggtgta 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 TATGGTGCTGATATCGAATCAAAAAACAAGCATGGCCTCACACCACTGTTACTTGGTGTA 840

Qy     841 catgagcaaaaaacagcaagtcgtgaaattttttaatcaagaaaaaaagcgaatttaaatgca 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 CATGAGCAAAAACAGCAAGTCGTGAAATTTTTAATCAAGAAAAAAGCGAATTTAAATGCA 900

Qy     901 ctggatagatatgtaaagactgctctctatacttgctgtatgttgtggatcagcaagtata 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 CTGGATAGATATGTAAGGACTGCTCTCATACTTGCTGTATGTTGTGGATCAGCAAGTATA 960

Qy     961 gtcagccttctacttgagcaaaatattgatgtatcttctcaagatctatctggacagacg 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 GTCAGCCTTCTACTTGAGCAAAATATTGATGTATCTTCTCAAGATCTATCTGGACAGACG 1020

Qy    1021 gccagagagtatgctgtttctagtcatcatcatgtaatttgccagttacttttctgactac 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 GCCAGAGAGTATGCTGTTTCTAGTCATCATCATGTAATTTGCCAGTTACTTTCTGACTAC 1080

Qy    1081 aaagaaaaacagatgctaaaaatctcttctgaaaacagcaatccagaaaatgtctcaa 1138
           |||||||||||||||||||||||||||||||||||||||||||||||||||| | | || ||
Db    1081 AAAGAAAAACAGATGCTAAAAATCTCTTCTGAAAACAGCAATCCAGAACAAGACTTAA 1138
```

RESULT 7
AX106588
LOCUS AX106588 1853 bp DNA PAT 30-APR-2001
DEFINITION Sequence 369 from Patent WO0125272.
ACCESSION AX106588
VERSION AX106588.1 GI:13922259
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1853)
AUTHORS Xu,J., Skeiky,Y.A., Reed,S.G. and Choever,M.A.
TITLE Compositions and methods for therapy and diagnosis of prostate cancer
JOURNAL Patent: WO 0125272-A 369 12-APR 2001;
CORIXA CORPORATION (US)
FEATURES Location/Qualifiers
source 1..1853
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 521 a 381 c 492 g 431 t 28 others
ORIGIN

Query Match 94.3%; Score 1088.6; DB 6; Length 1853;
Best Local Similarity 96.9%; Pred. No. 3.5e-266;

Matches 1104; Conservative 23, Mismatches 7, Indels 5, Gaps 2;

```
Qy       1 atggtggttgaggttgattccatgccggctgcctcttctgtgaagaagccatttggtctc 60
           ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db     223 ATGGTGGTTGAGGTTCATTCCATGCCGGCTGCCTCTTCTGTGAAGAAGCCATTTGGTCTC 282

Qy      61 aggagcaagatgggcaagtggtgctgccgttgcttcccctgctgcaggggagcggccaag 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     283 AGGAGCAAGATGGGCAAGTGGTGCTGCCGTTGCTTCCCCTGCTGCAGGGAGAGCGGCAAG 342

Qy     121 agcaacgtgggcacttctggagaccacgacgactctgctatgaaagactcagtaaacaag 180
           |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||| |||
Db     343 AGCAACGTGGGCACTTCTGGAGACCACGACGACTCTGCTATGAAGACACTCAGTAGCAAG 402

Qy     181 atgggcaaagtggtgccgccactgcttccccctgctgcagaggggagtgggcaagagcaacgtg 240
           |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db     403 ATGGGCAAGTGGTGCCGCCACTGCTTCCCCTGCTGCAGGGGAGTGGGCAAGAGCAACGTG 462

Qy     241 ggcgcttctggagaccacgacgactctgctatgaaagacactcaggaacaagatgggcaag 300
           |||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db     463 GGCGCTTCTGGAGACCACGACGAYTCTGCTATGAAGACACTCAGCAACAAGATGGGCAAG 522

Qy     301 tggtgctgccactgcttccccctgctgcaggggaggcggcaagagcaagtgggcgcttgg 360
           |||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db     523 TGGTGCTGCCACTGCTTCCCCTGCTGCAGGGGAGCRGCAAGAGCAAGTGGGCGCTTGG 582

Qy     361 ggagactacgatgacagtgccttcatggaggccaggtaccaactccgtggagaacatctg 420
           |||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db     583 GGAGACTACGATGACAGYGCCTTCATGGAKCCCAGGTACCACGTCCRTGGAGAACATCTG 642

Qy     421 gacaagctccacagagctgcctggtgggtaaagtccccagaaagaatctcatcgtcatg 480
           |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db     643 GACAAGCTCCACAGAGCTGCCTGGTGGGTAAAGTCCCCAGAAAGAATCTCATCGTCATG 702

Qy     481 ctcagggacactgacgtgaacaagaaggacaagcaaaagaggactgctctacatctggcc 540
           ||||||||||:||:||||||||||:|||||||||||||||||||||||||||||||||||||
Db     703 CTCAGGGACACKGAYGTGAACAAGAAGGACAAGCAAAAGAGGACTGCTCTACATCTGGCC 762

Qy     541 tctgccaatgggaattcagaagtagtaaaactcctgctggacagacgatgtcaacttaat 600
           ||||||||||||||||||||||||||||||||||||:|||||||||||||||||| |||||
Db     763 TCTGCCAATGGGAATTCAGAAGTAGTAAAACTCCTGCTGGACAGACGATGTCAACTTAAT 822

Qy     601 gtccttgacaacaaaaagaggacagctctgataaaggccgtacaatgccaggaagatgaa 660
           ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||||
Db     823 GTCCTTGACAACAAAAAGAGGACAGCTCTGAYAAAGGCCGTACAATGCCAGGAAGATGAA 882

Qy     661 tgtgcgttaatgttgctggaacatggcactgatccaaatattccagatgagtatggaaat 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     883 TGTGCGTTAATGTTGCTGGAACATGGCACTGATCCAAATATTCCAGATGAGTATGGAAAT 942

Qy     721 accactctgcactacgctatctataatgaaqataaattaatggccaaaqcactgctctta 780
           |||||||||.||||.|||.||||||||||||||||||||||||||||||||||||||||||.
Db     943 ACCACTCTRCACTAYGCTRTCTAYAATGAAGATAAATTAATGGCCAAAGCACTGCTCTTA 1002

Qy     781 tatggtgctgatatcgaatcaaaaaacaagcatggcctcacaccactgttacttggtgta 840
           ||:|||||||||||||||||||||||||||||||||||||||||||||||:||||||||:||
Db    1003 TAYGGTGCTGATATCGAATCAAAAAACAAGCATGGCCTCACACCACTGYTACTTGGTRTA 1062

Qy     841 catgagcaaaaaacagcaagtcgtgaaattttaatcaagaaaaaaagcgaatttaaatgc 899
           ||||||||||||||||||:|||||||||||||||:|||||||||||||||||||  |||||
Db    1063 CATGAGCAAAAACAGCAAGTSGTCAAATTTTTAATYAAGAAAAAAGCGAATTTAAAATGC 1122

Qy     900 actggatagatatgtaaagactgctctcatacttgctgtatgttgtggatcagcaagtat 959
           :|||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db    1123 RCTGGATAGATATGGAAGRACTGCTCTCATACTTGCTGTATGTTGTGGATCAGCAAGTAT 1182

Qy     960 agtcagccttctacttgagcaaaatattgatgtatcttctcaagatctatctggacagac 1019
           |||||||:|||||||||||||||:|||||||||||||||||||  |||||| ||||
Db    1183 AGTCAGCCYTCTACTTGAGCAAAATRTTGATGTATCTTCTCAAG----ATCTGGAAAGAC 1238

Qy     1020 ggccagagagtatgctgtttctagtcatcatcatgtaatttgccagttacttttctgacta 1079
           |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db     1239 GGCCAGAGAGTATGCTGTTTCTAGTCATCATCATGTAATTTGCCAGTTACTTTCTGACTA 1298
```

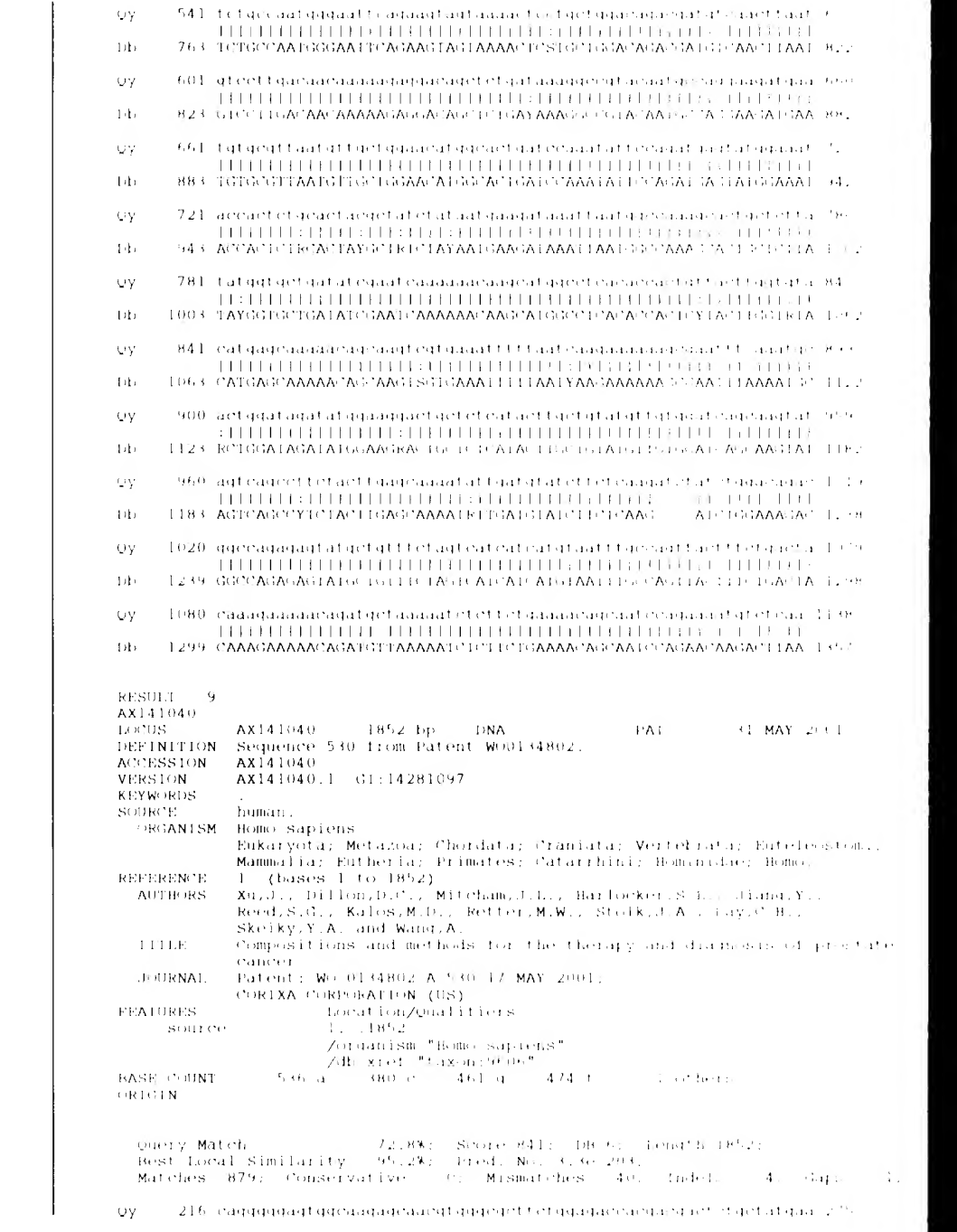


RESULT 8
AX140879
LOCUS AX140879 1853 bp DNA PAT 31-MAY-2001
DEFINITION Sequence 869 from Patent WO0134802.
ACCESSION AX140879
VERSION AX140879.1 GI:14280982
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1853)
AUTHORS Xu,J., Dillon,D.C., Mitcham,J.L., Barlocker,S.L., Jiang,Y., Reed,S.G., Kalos,M.D., Retter,M.W., Stolk,J.A., Day,C.H., Skolky,Y.A. and Wang,A.
TITLE Compositions and methods for the therapy and diagnosis of prostate cancer
JOURNAL Patent: WO 0134802 A 869 17-MAY-2001;
CORIXA CORPORATION (US)
FEATURES Location/Qualifiers
source 1..1853
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 521 a 381 c 492 g 431 t 28 others
ORIGIN

Query Match 94.8%; Score 1008.6; DB 6; Length 1853;
Best Local Similarity 96.2%; Pred. No. 3.5e-266;
Matches 1134; Conservative 23; Mismatches 7; Indels 5; Gaps 2;

RESULT 9
AX141040
LOCUS AX141040 1852 bp DNA PAT 31-MAY-2001
DEFINITION Sequence 530 from Patent WO0134802.
ACCESSION AX141040
VERSION AX141040.1 GI:14281097
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1852)
AUTHORS Xu,J., Dillon,D.C., Mitcham,J.L., Barlocker,S.L., Jiang,Y., Reed,S.G., Kalos,M.D., Retter,M.W., Stolk,J.A., Day,C.H., Skolky,Y.A. and Wang,A.
TITLE Compositions and methods for the therapy and diagnosis of prostate cancer
JOURNAL Patent: WO 0134802 A 530 17-MAY-2001;
CORIXA CORPORATION (US)
FEATURES Location/Qualifiers
source 1..1852
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 536 a 380 c 461 g 474 t 1 others
ORIGIN

Query Match 72.8%; Score 841; DB 6; Length 1852;
Best Local Similarity 95.2%; Pred. No. 3.6e-203;
Matches 879; Conservative 0; Mismatches 40; Indels 4; Gaps 1;



```
          ||||||||| |||||||||||||||||| |||||||||||||| |||||||  || ||||
Db    771 CAGGGGGAGCGGCAAGAGCAACGTGGGCACTTCTGGAGACCACAACGACTCCTCTGTGAA 830

Qy    276 gacactcaggaacaagatgggcaagtggtgctgccactgcttcccctgctgcagggggag 335
          ||| ||   ||| ||||| | |||||||||||||||||||||||||||||||||||||||
Db    831 GACGCTTGGGAGCAAGAGGTGCAAGTGGTGCTGCCACTGCTTCCCCTGCTGCAGGGGCAG 890

Qy    336 cggcaagagcaaggtgggcgcttggggagactacgatgacagtgccttcatggagcccag 395
          |||||||||||| |||| |||||||||||||||||||||||| ||||||||||  |||||
Db    891 CGGCAAGAGCAACGTGGTCGCTTGGGGAGACTACGATGACAGCCCCTTCATGGATCCCAG 950

Qy    396 gtaccacgtccgtggagaagatctggacaagctccacagagctgcctggtggggtaaagt 455
          ||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db    951 GTACCACGTCCATGGAGAAGATCTGGACAAGCTCCACAGAGCTGCCTGGTGGGGTAAAGT 1010

Qy    456 ccccagaaaggatctcatcgtcatgctcagggacactgacgtgaacaagaaggacaagca 515
          |||||||||||||||||||||||||||||||||||| || |||||||||| |||||||||
Db   1011 CCCCAGAAAGGATCTCATCGTCATGCTCAGGGACACGGATGTGAACAAGAGGGACAAGCA 1070

Qy    516 aaagaggactgctctacatctggcctctgccaatgggaattcagaagtagtaaaactcct 575
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db   1071 AAAGAGGACTGCTCTACATCTGGCCTCTGCCAATGGGAATTCAGAAGTAGTAAAACTCGT 1130

Qy    576 gctggacagacgatgtcaacttaatgtccttgacaacaaaaagaggacagctctgataaa 635
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db   1131 GCTGGACAGACGATGTCAACTTAATGTCCTTGACAACAAAAAGAGGACAGCTCTGACAAA 1190

Qy    636 ggccgtacaatgccaggaagatgaatgtgcgttaatgttgctggaacatggcactgatcc 695
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1191 GGCCGTACAATGCCAGGAAGATGAATGTGCGTTAATGTTGCTGGAACATGGCACTGATCC 1250

Qy    696 aaatattccagatgagtatggaaataccactctgcactacgctatctataatgaagataa 755
          ||||||||||||||||||||||||||||||||| ||||| ||| |||| |||||||||||
Db   1251 AAATATTCCAGATGAGTATGGAAATACCACTCTACACTATGCTGTCTACAATGAAGATAA 1310

Qy    756 attaatggccaaagcactgctcttatatggtgctgatatcgaatcaaaaaacaagcatgg 815
          ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db   1311 ATTAATGGCCAAAGCACTGCTCTTATACGGTGCTGATATCGAATCAAAAAACAAGCATGG 1370

Qy    816 cctcacaccactgttacttggtgtacatgagcaaaaacagcaagtcgtgaaatttttaat 875
          ||||||||||||| |||||||| |||||||||||||||||||||| ||||||||||||||
Db   1371 CCTCACACCACTGCTACTTGGTATACATGAGCAAAAACAGCAAGTGGTGAAATTTTTAAT 1430

Qy    876 caagaaaaaagcgaatttaaatgcactggatagatatggaaggactgctctcatacttgc 935
          |||||||||||||||||||||||| |||||||||||||||||| ||||||||||||||||
Db   1431 CAAGAAAAAAGCGAATTTAAATGCGCTGGATAGATATGGAAGAACTGCTCTCATACTTGC 1490

Qy    936 tgtatgttgtggatcagcaagtatagtcagccttctacttgagcaaaatattgatgtatc 995
          |||||||||||||||||||||||||||||||| |||||||||||||||| ||||||||||
Db   1491 TGTATGTTGTGGATCAGCAAGTATAGTCAGCCCTCTACTTGAGCAAAATGTTGATGTATC 1550

Qy    996 ttctcaagatctatctggacagacggccagagagtatgctgtttctagtcatcatcatgt 1055
          ||||||||    ||||||| ||||||||||||||||||||||||||||||||||||||||
Db   1551 TTCTCAAG----ATCTGGAAAGACGGCCAGAGAGTATGCTGTTTCTAGTCATCATCATGT 1606

Qy   1056 aatttgccagttactttctgactacaaagaaaaacagatgctaaaaatctcttctgaaaa 1115
          |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db   1607 AATTTGCCAGTTACTTTCTGACTACAAAGAAAAACAGATGTTAAAAATCTCTTCTGAAAA 1666

Qy   1116 cagcaatccagaaaatgtctcaa 1138
          ||||||||||||| | | || ||
Db   1667 CAGCAATCCAGAACAAGACTTAA 1689


RESULT 10
AR148112/c
LOCUS       AR148112     1851 bp    DNA             PAT       08-AUG-2001
DEFINITION  Sequence 291 from patent US 6225054.
ACCESSION   AR148112
VERSION     AR148112.1  GI:15112202
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1851)
  AUTHORS   Frudakis,T.N., Smith,J.M. and Reed,S.G.
  TITLE     Compositions and methods for the treatment and diagnosis of breast
            cancer
  JOURNAL   Patent: US 6225054-A 291 01-MAY-2001;
FEATURES             Location/Qualifiers
     source          1..1851
                     /organism="unknown"
BASE COUNT      474 a    460 c    380 g    536 t      1 others
ORIGIN


  Query Match             71.8%;  Score 929;  DB 6;  Length 1851;
  Best Local Similarity   95.1%;  Pred. No. 3.7e-200;
  Matches  878;  Conservative    0;  Mismatches   40;  Indels    5;  Gaps    2;

Qy    216 cagggggagtggcaagagcaacgtgggcgcttctggagaccacgacgactctgctatgaa 275
          ||||||||| |||||||||||||||||| |||||||||||||| |||||||  || ||||
Db   1081 CAGGGGGAGCGGCAAGAGCAACGTGGGCACTTCTGGAGACCACAACGACTCCTCTGTGAA 1022

Qy    276 gacactcaggaacaagatgggcaagtggtgctgccactgcttcccctgctgcagggggag 335
          ||| ||   ||| ||||| | |||||||||||||||||||||||||||||||| ||||||
Db   1021 GACGCTTGGGAGCAAGAGGTGCAAGTGGTGCTGCCACTGCTTCCCCTGCTGCA-GGGGAG 963

Qy    336 cggcaagagcaaggtgggcgcttggggagactacgatgacagtgccttcatggagcccag 395
          |||||||||||| |||| |||||||||||||||||||||||| ||||||||||  |||||
Db    962 CGGCAAGAGCAACGTGGTCGCTTGGGGAGACTACGATGACAGCGCCTTCATGGATCCCAG 903

Qy    396 gtaccacgtccgtggagaagatctggacaagctccacagagctgcctggtggggtaaagt 455
          ||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db    902 GTACCACGTCCATGGAGAAGATCTGGACAAGCTCCACAGAGCTGCCTGGTGGGGTAAAGT 843

Qy    456 ccccagaaaggatctcatcgtcatgctcagggacactgacgtgaacaagaaggacaagca 515
          |||||||||||||||||||||||||||||||||||| || |||||||||| |||||||||
Db    842 CCCCAGAAAGGATCTCATCGTCATGCTCAGGGACACGGATGTGAACAAGAGGGACAAGCA 783

Qy    516 aaagaggactgctctacatctggcctctgccaatgggaattcagaagtagtaaaactcct 575
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db    782 AAAGAGGACTGCTCTACATCTGGCCTCTGCCAATGGGAATTCAGAAGTAGTAAAACTCGT 723

Qy    576 gctggacagacgatgtcaacttaatgtccttgacaacaaaaagaggacagctctgataaa 635
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db    722 GCTGGACAGACGATGTCAACTTAATGTCCTTGACAACAAAAAGAGGACAGCTCTGACAAA 663

Qy    636 ggccgtacaatgccaggaagatgaatgtgcgttaatgttgctggaacatggcactgatcc 695
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    662 GGCCGTACAATGCCAGGAAGATGAATGTGCGTTAATGTTGCTGGAACATGGCACTGATCC 603

Qy    696 aaatattccagatgagtatggaaataccactctgcactacgctatctataatgaagataa 755
          ||||||||||||||||||||||||||||||||| ||||| ||| |||| |||||||||||
Db    602 AAATATTCCAGATGAGTATGGAAATACCACTCTACACTATGCTGTCTACAATGAAGATAA 543

Qy    756 attaatggccaaagcactgctcttatatggtgctgatatcgaatcaaaaaacaagcatgg 815
          ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db    542 ATTAATGGCCAAAGCACTGCTCTTATACGGTGCTGATATCGAATCAAAAAACAAGCATGG 483

Qy    816 cctcacaccactgttacttggtgtacatgagcaaaaacagcaagtcgtgaaatttttaat 875
          ||||||||||||| |||||||| |||||||||||||||||||||| ||||||||||||||
Db    482 CCTCACACCACTGCTACTTGGTATACATGAGCAAAAACAGCAAGTGGTGAAATTTTTAAT 423

Qy    876 caagaaaaaagcgaatttaaatgcactggatagatatggaaggactgctctcatacttgc 935
          |||||||||||||||||||||||| |||||||||||||||||| ||||||||||||||||
Db    422 CAAGAAAAAAGCGAATTTAAATGCGCTGGATAGATATGGAAGAACTGCTCTCATACTTGC 363

Qy    936 tgtatgttgtggatcagcaagtatagtcagccttctacttgagcaaaatattgatgtatc 995
          |||||||||||||||||||||||||||||||| |||||||||||||||| ||||||||||
Db    362 TGTATGTTGTGGATCAGCAAGTATAGTCAGCCCTCTACTTGAGCAAAATGTTGATGTATC 303

Qy    996 ttctcaagatctatctggacagacggccagagagtatgctgtttctagtcatcatcatgt 1055
          ||||||||    ||||||| ||||||||||||||||||||||||||||||||||||||||
Db    302 TTCTCAAG----ATCTGGAAAGACGGCCAGAGAGTATGCTGTTTCTAGTCATCATCATGT 247
```

```
Qy       1056 aatttgccagttactttctgactacaaagaaaaacagatgctaaaaatctcttctgaaaa 1115
Db        246 AATTTGCCAGTTACTTTCTGACTACAAAGAAAAACAGATGTTAAAAATCTCTTCTGAAAA 187

Qy       1116 cagcaatccagaaaatgtctcaa 1138
Db        186 CAGCAATCCAGAACAAGACTTAA 164


RESULT 11
AX106585/c
LOCUS       AX106585    1851 bp    DNA             PAT       30-APR-2001
DEFINITION  Sequence 366 from Patent WO0125272.
ACCESSION   AX106585
VERSION     AX106585.1  GI:13922256
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1851)
  AUTHORS   Xu,J., Skeiky,Y.A., Reed,S.G. and Cheever,M.A.
  TITLE     Compositions and methods for therapy and diagnosis of prostate
            cancer
  JOURNAL   Patent: WO 0125272-A 366 12-APR-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..1851
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      474 a    460 c    380 g    536 t      1 others
ORIGIN

  Query Match             71.8%;  Score 829;  DB 6;  Length 1851;
  Best Local Similarity   95.1%;  Pred. No. 3.7e-200;
  Matches  878;  Conservative    0;  Mismatches   40;  Indels    5;  Gaps    2;

Qy        216 caggggagtggcaagagcaacgtgggcgcttctggagaccacgacgactctgctatgaa 275
Db       1081 CAGGGGAGCGGCAAGAGCAACGTGGGCACTTCTGGAGACCACAACGACTCCTCTGTGAA 1022

Qy        276 gacactcaggaacaagatgggcaagtggtgctgccactgcttcccctgctgcaggggag 335
Db       1021 GACGCTTGGGAGCAAGAGGTGCAAGTGGTGCTGCCACTGCTTCCCCTGCTGCA-GGGGAG 963

Qy        336 cggcaagagcaaagtgggcgcttggggagactacgatgacagtgccttcatggagcccag 395
Db        962 CGGCAAGAGCAACGTGGTCGCTTGGGGAGACTACGATGACAGCGCCTTCATGGATCCCAG 903

Qy        396 gtaccacgtccatggagaagatctggacaagctccacagagctgcctggtggggtaaagt 455
Db        902 GTACCACGTCCATGGACAAGATCTGGACAAGCTCCACAGAGCTGCCTGGTGGGGTAAAGT 843

Qy        456 ccccagaaaggatctcatcgtcatgctcagggacactgacgtgaacaagaaggacaagca 515
Db        842 CCCCAGAAAGGATCTCATCGTCATGCTCAGGGACACGGATGTGAACAAGAGGGACAAGCA 783

Qy        516 aaagaggactgctctacatctggcctctgccaatgggaattcagaagtagtaaaactcct 575
Db        782 AAAGAGGACTGCTCTACATCTGGCCTCTGCCAATGGGAATTCAGAAGTAGTAAAACTCGT 723

Qy        576 gctggacagacgatgtcaacttaatgtccttgacaacaaaaagaggacagctctgataaa 635
Db        722 GCTGGACAGACGATGTCAACTTAATGTCCTTGACAACAAAAAGAGGACAGCTCTGACAAA 663

Qy        636 ggccgtacaatgccaggaagatgaatgtgcgttaatgttgctggaacatggcactgatcc 695
Db        662 GGCCGTACAATGCCAGGAAGATGAATGTGCGTTAATGTTGCTGGAACATGGCACTGATCC 603

Qy        696 aaatattccagatgagtatggaaatacccctctgcactacgctatctataatgaagataa 755
Db        602 AAATATTCCAGATGAGTATGGAAATACCACTCTACACTATGCTGTCTACAATGAAGATAA 543

Qy        756 attaatggccaaagcactgctcttatatggtgctgatatcgaatcaaaaaacaagcatgg 815
Db        542 ATTAATGGCCAAAGCACTGCTCTTATACGGTGCTGATATCGAATCAAAAAACAAGCATGG 483

Qy        816 cctcacaccactgttacttggtgtacatgagcaaaaacagcaagtcgtgaaatttttaat 875
Db        482 CCTCACACCACTGCTACTTGGTATACATGAGCAAAAACAGCAAGTGGTGAAATTTTTAAT 423

Qy        876 caagaaaaaagcgaatttaaatgcactggatagatatggaagaactgctctcatacttgc 935
Db        422 CAAGAAAAAAGCGAATTTAAATGCGCTGGATAGATATGGAAGAACTGCTCTCATACTTGC 363

Qy        936 tgtatgttgtggatcagcaagtatagtcagccttctacttgagcaaaatattgatgtatc 995
Db        362 TGTATGTTGTGGATCAGCAAGTATAGTCAGCCTTCTACTTGAGCAAAATGTTGATGTATC 303

Qy        996 ttctcaagatctatctggacagacggccagagagtatgctgtttctagtcatcatcatgt 1055
Db        302 TTCTCAAG     ATCTGGACAGACGGCCAGAGAGTATGCTGTTTCTAGTCATCATCATGT 247

Qy       1056 aatttgccagttactttctgactacaaagaaaaacagatgctaaaaatctcttctgaaaa 1115
Db        246 AATTTGCCAGTTACTTTCTGACTACAAAGAAAAACAGATGTTAAAAATCTCTTCTGAAAA 187

Qy       1116 cagcaatccagaaaatgtctcaa 1138
Db        186 CAGCAATCCAGAACAAGACTTAA 164


RESULT 12
AX140876/c
LOCUS       AX140876    1851 bp    DNA             PAT       31-MAY-2001
DEFINITION  Sequence 366 from Patent WO0134802.
ACCESSION   AX140876
VERSION     AX140876.1  GI:14280979
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1851)
  AUTHORS   Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,
            Reed,S.G., Kalos,M.D., Fetter,M.W., Stolk,J.A., Day,C.H.,
            Skeiky,Y.A. and Wang,A.
  TITLE     Compositions and methods for the therapy and diagnosis of prostate
            cancer
  JOURNAL   Patent: WO 0134802-A 366 17-MAY-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..1851
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      474 a    460 c    380 g    536 t      1 others
ORIGIN

  Query Match             71.8%;  Score 829;  DB 6;  Length 1851;
  Best Local Similarity   95.1%;  Pred. No. 3.7e-200;
  Matches  878;  Conservative    0;  Mismatches   40;  Indels    5;  Gaps    2;

Qy        216 caggggagtggcaagagcaacgtgggcgcttctggagaccacgacgactctgctatgaa 275
Db       1081 CAGGGGAGCGGCAAGAGCAACGTGGGCACTTCTGGAGACCACAACGACTCCTCTGTGAA 1022

Qy        276 gacactcaggaacaagatgggcaagtggtgctgccactgcttcccctgctgcaggggag 335
Db       1021 GACGCTTGGGAGCAAGAGGTGCAAGTGGTGCTGCCACTGCTTCCCCTGCTGCA-GGGGAG 963

Qy        336 cggcaagagcaaagtgggcgcttggggagactacgatgacagtgccttcatggagcccag 395
Db        962 CGGCAAGAGCAACGTGGTCGCTTGGGGAGACTACGATGACAGCGCCTTCATGGATCCCAG 903
```

```
Qy     396 gtaccacgtccgtggaaaagatctggacaagctccacagagctgcctggtgggggtaaaagt 455
           |||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     902 GTACCACGTCCATGGAGAAGATCTGGACAAGCTCCACAGAGCTGCCTGGTGGGGTAAAGT 843

Qy     456 ccccagaaaggatctcatcgtcatgctcagggacactgacgtgaacaagaaggacaagca 515
           |||||||||||||||||||||||||||||||||||| || |||||||||| |||||||||
Db     842 CCCCAGAAAGGATCTCATCGTCATGCTCAGGGACACGGATGTGAACAAGAGGGACAAGCA 783

Qy     516 aaagaggactgctctacatctggcctctgccaatgggaattcagaagtagtaaaactcct 575
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db     782 AAAGAGGACTGCTCTACATCTGGCCTCTGCCAATGGGAATTCACAACTAGTAAAACTCGT 723

Qy     576 gctggacagacgatgtcaacttaatgtccttgacaacaaaaagagggacagctctgataaa 635
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db     722 GCTGGACAGACGATGTCAACTTAATGTCCTTGACAACAAAAAGAGGACAGCTCTGACAAA 663

Qy     636 ggccgtacaatgccaggaagatgaatgtgcgttaatgttgctggaacatggcactgatcc 695
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     662 GGCCGTACAATGCCAGGAAGATGAATGTGCGTTAATGTTGCTGGAACATGGCACTGATCC 603

Qy     696 aaatattccagatgagtatggaaatacccactctgcactacgctatctataatgaagataa 755
           |||||||||||||||||||||||||||||||||| |||| ||| |||| ||||||||||
Db     602 AAATATTCCAGATGAGTATGGAAATACCACTCTACACTATGCTGTCTACAATGAAGATAA 543

Qy     756 attaatggccaaagcactgctcttatatgctgctgatatcgaatcaaaaaacaagcatgg 815
           |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db     542 ATTAATGGCCAAAGCACTGCTCTTATACGGTGCTGATATCGAATCAAAAAACAAGCATGG 483

Qy     816 cctcacaccactgttacttggtgtacatgaagcaaaaacagcaagtcgtgaaatttttaat 875
           |||||||||||| |||||||| |||||||||||||||||||| ||||||||||||||||
Db     482 CCTCACACCACTGCTACTTGGTATACATGAGCAAAAACAGCAAGTGGTGAAATTTTTAAT 423

Qy     876 caagaaaaaaagcgaatttaaatgcactggatagatatgaaagcactgctctcatacttgc 935
           |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db     422 CAAGAAAAAAGCGAATTTAAATGCACTGGATAGATATGGAAGACTGCTCTCATACTTGC 363

Qy     936 tgtatgttgtggatcagcaagtatagtcagcctctacttgagcaaaatgttgatgtatc 995
           |||||||||||||||||||||||||||||||| ||||||||||||||| ||||||||||
Db     362 TGTATGTTGTGGATCAGCAAGTATAGTCAGCCCTCTACTTGAGCAAAATGTTGATGTATC 303

Qy     996 ttctcaagatctatctggacagacggccagagagtatgctgtttctagtcatcatcatgt 1055
           |||||||    |||||| |||||||||||||||||||||||||||||||||||||||||
Db     302 TTCTCAAG----ATCTGGAAAGACGGCCAGAGAGTATGCTGTTTCTAGTCATCATCATGT 247

Qy    1056 aatttgccagttacttttctgactacaaagaaaaacagatgctaaaaatctcttctgaaaa 1115
           |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db     246 AATTTGCCAGTTACTTTCTGACTACAAAGAAAAACAGATGTTAAAAATCTCTTCTGAAAA 187

Qy    1116 caqcaatccagaaaatgtctcaa 1138
           ||||||||||||| | | || ||
Db     186 CAGCAATCCAGAACAAGACTTAA 164


RESULT  13
AX106589
LOCUS       AX106589     2184 bp    DNA             PAT       30-APR-2001
DEFINITION  Sequence 370 from Patent WO0125272.
ACCESSION   AX106589
VERSION     AX106589.1  GI:13922260
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi,
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2184)
  AUTHORS   Xu,J., Skeiky,Y.A., Reed,S.G. and Cheever,M.A.
  TITLE     Compositions and methods for therapy and diagnosis of prostate
            cancer
  JOURNAL   Patent: WO 0125272-A 370 12-APR-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..2184
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      635 a     447 c     536 g     563 t       3 others
ORIGIN


  Query Match             71.8%;  Score 829;  DB 6;  Length 2184;
  Best Local Similarity   95.1%;  Pred. No. 2.7e-209;
  Matches  878;  Conservative    0,  Mismatches   40;  Indels    5;  Gaps    2;

Qy     216 cagqggqaqtggcaagaqcaacgtqqqcgcttctqqaqaccacgacgactctgctatgaa 275
           |||||||| |||||||||||||||| |||||||||| |||||||  || ||||
Db     771 CAGGGGCAGCGGCAAGAGCAACGTGGGCACTTCTGGAGACCACAACGACTCCTCTGTGAA 830

Qy     276 gacactcaqqaacaaqatqqqcaaatqqtgctqccactqcttcccctgctqccaqqqtaa 335
           ||| ||  ||| ||||| | |||||||||||||||||||||||||||||||  |||||
Db     831 GACGCTTGGGAGCAAGAGCTCCAAGTGGTGCTGCCACTGCTTCCCCTGCTGCA-GGGGAG 889

Qy     336 cqqcaaqaqcaaqgtgqqcgcttqqqqagactacqatgacaqtgccttcatqqaacccaq 395
           |||||||||||| |||| |||||||||||||||||||||||||| ||||||||||| |||||
Db     890 CGGCAAGAGCAACGTGGTCGCTTGGGGAGACTACGATGACAGCGCCTTCATGGAACCCAG 949

Qy     396 gtaccacgtccgtggaqaaqatctqqacaaqctccacaqaqctqcctggtqqqqtaaaqt 455
           |||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db     950 GTACCACGTCCATGGAGAAGATCTGGACAAGCTCCACAGAGCTGCCTGGTGGGGTAAAGT 1009

Qy     456 ccccagaaaggatctcatcgtcatgctcagggacactgacgtgaacaagaaggacaagca 515
           |||||||||||||||||||||||||||||||| || |||||||||| |||||||||
Db    1010 CCCCAGAAAGGATCTCATCGTCATGCTCAGGGACACGGATGTGAACAAGAGGGACAAGCA 1069

Qy     516 aaagaggactgctctacatctggcctctgccaatgggaattcagaagtagtaaaactcct 575
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db    1070 AAAGAGGACTGCTCTACATCTGGCCTCTGCCAATGGGAATTCAGAAGTAGTAAAACTCGT 1129

Qy     576 gctggacagacgatgtcaacttaatgtccttgacaacaaaaagagggacagctctgataaa 635
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db    1130 GCTGGACAGACGATGTCAACTTAATGTCCTTGACAACAAAAAGAGGACAGCTCTGACAAA 1189

Qy     636 ggccgtacaatgccaggaagatgaatgtgcgttaatgttgctggaacatggcactgatcc 695
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1190 GGCCGTACAATGCCAGGAACATGAATGTGCGTTAATGTTGCTGGAACATGGCACTGATCC 1249

Qy     696 aaatattccagatgagtatggaaatacccactctgcactacgctatctataatgaagataa 755
           |||||||||||||||||||||||||||||||| |||| ||| |||| ||||||||||
Db    1250 AAATATTCCAGATGAGTATGGAAATACCACTCTACACTATGCTGTCTACAATGAAGATAA 1309

Qy     756 attaatggccaaagcactgctcttatatgctgctgatatcgaatcaaaaaacaagcatgg 815
           |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db    1310 ATTAATGGCCAAAGCACTGCTCTTATACGGTGCTGATATCGAATCAAAAAACAAGCATGG 1369

Qy     816 cctcacaccactgttacttggtgtacatgaagcaaaaacagcaagtcgtgaaatttttaat 875
           |||||||||||| |||||||| |||||||||||||||||||| ||||||||||||||||
Db    1370 CCTCACACCACTGCTACTTGGTATACATGAGCAAAAACAGCAAGTGGTGAAATTTTTAAT 1429

Qy     876 caagaaaaaaagcgaatttaaatgcactggatagatatgaaagcactgctctcatacttgc 935
           |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db    1430 CAAGAAAAAAGCGAATTTAAATGCACTGGATAGATATGGAAGACTGCTCTCATACTTGC 1489

Qy     936 tgtatgttgtggatcagcaagtatagtcagcctctacttgagcaaaatattgatgtatc 995
           |||||||||||||||||||||||||||||||| |||||||||||||||| |||| ||||
Db    1490 TGTATGTTGTGGATCAGCAAGTATAGTCAGCCCTCTACTTGAGCAAAATGTTGATGTATC 1549

Qy     996 ttctcaagatctatctggacagacggccagagagtatgctgtttctagtcatcatcatgt 1055
           |||||||    |||||| |||||||||||||||||||||||||||||||||||||||||
Db    1550 TTCTCAAG----ATCTGGAAAGACGGCCAGAGAGTATGCTGTTTCTAGTCATCATCATGT 1605

Qy    1056 aatttgccagttacttttctgactacaaagaaaaacagatgctaaaaatctcttctgaaaa 1115
           |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db    1606 AATTTGCCAGTTACTTTCTGACTACAAAGAAAAACAGATGTTAAAAATCTCTTCTGAAAA 1665

Qy    1116 caqcaatccagaaaatgtctcaa 1138
           ||||||||||||| | | || ||
Db    1666 CAGCAATCCAGAACAAGACTTAA 1688
```

```
RESULT 14
AX140880
LOCUS       AX140880     2184 bp    DNA             PAT       31-MAY-2001
DEFINITION  Sequence 370 from Patent WO0134802.
ACCESSION   AX140880
VERSION     AX140880.1  GI:14280983
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2184)
  AUTHORS   Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,
            Reed,S.G., Kalos,M.D., Fetter,M.W., Stolk,J.A., Day,C.H.,
            Skeiky,Y.A. and Wang,A.
  TITLE     Compositions and methods for the therapy and diagnosis of prostate
            cancer
  JOURNAL   Patent: WO 0134802-A 370 17-MAY-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..2184
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      635 a    447 c    536 g    563 t      3 others
ORIGIN

  Query Match             71.8%;  Score 828;  DB 6;  Length 2184;
  Best Local Similarity   95.1%;  Pred. No. 3.7e-200;
  Matches  878;  Conservative    0;  Mismatches   40;  Indels    5;  Gaps    2;

Qy        216 ctggttgaggtggcaagagcaacgtgggcgcttctggagaccacgacgactctgctatgaa 275
Db        771 CAGGTGGAGGGCAAGAGCAACGTGGGCACTTCTGGAGACCACAACGACTCCTCTGTGAA 830

Qy        276 ggacctcaggaacaagatgggcaagtggtgctgccactgcttcccctgctgcagggggag 335
Db        831 GACCTTGGGAGCAAGAGGTGCAAGTGGTGCTGCCACTGCTTCCCCTGCTGCA GGGGAG 889

Qy        336 cggctaatgcaagctggcggcttggggagactacgatgacagtgccttcatggagcccag 395
Db        890 CGGCAACAGCAACGTGGTCGGCTTGGGGAGACTACGATGACAGCGCCTTCATGGATCCCAG 949

Qy        396 gtacccagtcagtggagaagatctggacaagctccacagagctgcctggtggggtaaagt 455
Db        950 GTACCACGTCCATGGAGAAGATCTGGACAAGCTCCACAGAGCTGCCTGGTGGGGTAAAGT 1009

Qy        456 cccccagaaaggatctcatcgtcatgctcagggacactgacgtgaacaagaaggacaagca 515
Db       1010 CCCCAGAAAGGATCTCATCGTCATGCTCAGGGACACTGGATGTGAACAAGAGGGACAAGCA 1069

Qy        516 aaagaggactgctctacatctggcctctgccaatggaaattcagaagtagtaaaactcct 575
Db       1070 AAAGAGGACTGCTCTACATCTGGCCTCTGCCAATGGAATTCAGAAGTAGTAAAACTCGT 1129

Qy        576 gctggacagacgatgtcaacttaatgtccttgacaacaaaaagaggacagctctgataaa 635
Db       1130 GCTGGACAGACGATGTCAACTTAATGTCCTTGACAACAAAAAGAGGACAGCTCTGACAAA 1189

Qy        636 ggccgtacaatgccaggaagatgaatgtgcgttaatgttgctggaacatggcactgatcc 695
Db       1190 GGCCGTACAATGCCAGGAAGATGAATGTGCGTTAATGTTGCTGGAACATGGCACTGATCC 1249

Qy        696 aaatattccagatgagtatggaaataccactctgcactacgctatctataatgaagataa 755
Db       1250 AAATATTCCAGATGAGTATGGAAATACCACTCTACACTATGCTGTCTACAATGAAGATAA 1309

Qy        756 attaatggccaaagcactgctcttatatggtgctgatatcgaatcaaaaaacaagcatgg 815
Db       1310 ATTAATGGCCAAAGCACTGCTCTTATACGGTGCTGATATCGAATCAAAAAACAAGCATGG 1369

Qy        816 cctcacaccactgttacttggtgtacatgagcaaaaacagcaagtcgtgaaatttttaat 875
Db       1370 CCTCACACCACTGCTACTTGGTATACATGAGCAAAAACAGCAAGTGGTGAAATTTTTAAT 1429

Qy        876 caagaaaaagcgaatttaaatgcactggatagatatggaaggactgctctcatacttgc 935
Db       1430 CAAGAAAAAGCGAATTTAAATGCGCTGGATAGATATGGAAGAACTGCTCTCATACTTGC 1489

Qy        936 tgtatgttgtggatcagcaagtatagtcagccttctacttgagcaaaatattgatgtatc 995
Db       1490 TGTATGTTGTGGATCAGCAAGTATAGTCAGCCTTCTACTTGAGCAAAATGTTGATGTATC 1549

Qy        996 ttctcaagatctatctggacagacggccagagagtatgctgtttctagtcatcatcatgt [illegible]
Db       1550 TTCTCAAG     ATCTGGAAAGACGGCCAGAGAGTATGCTGTTTCTAGTCATCATCATGT [illegible]

Qy       1056 aatttgccagttactttctgactacaaagaaaacagatgctaaaaatctgtttctgaaaa [illegible]
Db       1606 AATTTGCCAGTTACTTTCTGACTACAAAGAAAACAGATGTTAAAAATCTTTTCTGAAAA [illegible]

Qy       1116 cagcaatccagaaaatgtctcaa 1138
Db       1666 CAGCAATCCAGAACAAGACTTAA 1688

RESULT 15
AX106587
LOCUS       AX106587     1512 bp    DNA             PAT       30-APR-2001
DEFINITION  Sequence 368 from Patent WO0125272.
ACCESSION   AX106587
VERSION     AX106587.1  GI:13922258
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1512)
  AUTHORS   Xu,J., Skeiky,Y.A., Reed,S.G. and Cheever,M.A.
  TITLE     Compositions and methods for therapy and diagnosis of prostate
            cancer
  JOURNAL   Patent: WO 0125272-A 368 12-APR-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..1512
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      406 a    301 c    393 g    399 t     13 others
ORIGIN

  Query Match             69.7%;  Score 804.8;  DB 6;  Length 1512;
  Best Local Similarity   98.8%;  Pred. No. 5.3e-194;
  Matches  800;  Conservative    9;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 atggtggttgaggttgattccatgccggctgccttcttctgtgaagaagcatttggtctc 60
Db        223 ATGGTGGTTGAGGTTGATTCCATGCCGGCTGCTTCTTCTGTGAAGAAGCATTTGGTCTC 282

Qy         61 aggagcaagatgggcaagtggtgctgccgttgcttcccctgctgcagggagagcggcaag 120
Db        283 AGGAGCAAGATGGGCAAGTGGTGCTGCCGTTGCTTCCCCTGCTGCAGGGAGAGCGGCAAG 342

Qy        121 agcaacgtgggcacttctggagaccacgacgactctgctatgaagacactcaggagcaag 180
Db        343 AGCAACGTGGGCACTTCTGGAGACCACGACGACTCTGCTATGAAGACACTCAGGAGCAAG 402

Qy        181 atgggcaagtggtgccgccactgcttcccctgctgcaggggagtggcaagagcaacgtg 240
Db        403 ATGGGCAAGTGGTGCCGCCACTGCTTCCCCTGCTGCAGGGGAGTGGCAAGAGCAACGTG 462

Qy        241 ggcgcttctggagaccacgacgactctgctatgaagacactcaggaacaagatgggcaag 300
Db        463 GGCGCTTCTGGAGACCACGACGAYTCTGCTATGAAGACACTCAGGAACAAGATGGGCAAG 522
```

```
Qy      301 tqqtqctqccactqcttcccctqctqcaqqqqqaqcqqcaaqaqcaaqqtqqqcqcttqq 360
            ||||||||||||||||||||||||||||||||||||||:||||||||||||||||||||||
Db      523 TGGTGCTGCCACTGCTTCCCCTGCTGCAGGGGGAGCRGCAAGAGCAAGGTGGGCGCTTGG 582

Qy      361 qqaqactacqatqacaqtqccttcatqqaqcccaqqtaccacqtccqtqgaqaaqatctq 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      583 GGAGACTACGATGACAGTGCCTTCATGGAGCCCAGGTACCACGTCCGTGGAGAAGATCTG 642

Qy      421 qacaaqctccacacqaqctqccctqqtqqqqtqaaqtccccaqaaaqqatctcatcqtcatq 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      643 GACAAGCTCCACAGAGCTGCCTGGTGGGGTAAAGTCCCCAGAAAGGATCTCATCGTCATG 702

Qy      481 ctcaqqqacactqacqtqaaacaaqaaqqacaaqcaaaaqaqqactqctctacatctqqcc 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      703 CTCAGGGACACTGACGTGAACAAGAAGGACAAGCAAAAGAGGACTGCTCTACATCTGGCC 762

Qy      541 tctqccaatqqqaattcaqaaqtaqtaaaactcctqctqqacaqacqatqtcaacttaat 600
            ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||
Db      763 TCTGCCAATGGGAATTCAGAAGTAGTAAAACTCSTGCTGGACAGACGATGTCAACTTAAT 822

Qy      601 qtccttqacaacaaaaaqaqqacaqctctqataaaqqccqtacaatqccaqqaaqatqaa 660
            ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||
Db      823 GTCCTTGACAACAAAAAGAGGACAGCTCTGAYAAAGGCCGTACAATGCCAGGAAGATGAA 882

Qy      661 tqtqcqttaatqttqctqqaacatqqcactqatccaaatattccaqatqaqtatqqaaat 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      883 TGTGCGTTAATGTTGCTGGAACATGGCACTGATCCAAATATTCCACATGAGTATGGAAAT 942

Qy      721 accactctqcactacqctatctataatqaaqataaattaatqqccaaaqcactqctctta 780
            |||||||:||||:|||:||||:||||||||||||||||||||||||||||||||||||
Db      943 ACCACICTRCACTAYGCTRTCTAYAATGAAGATAAATTAATGGCCAAAGCACTGCTCTTA 1002

Qy      781 tatqqtqctqatatcqaatcaaaaaacaaq 810
            ||:|||||||||||||||||||||||||||
Db     1003 TAYGGTGCTGATATCGAATCAAAAAACAAG 1032
```

Search completed: March 9, 2002, 02:19:48
Job time: 6507 sec

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: March 9, 2002, 02:19:52 ; Search time 1487.31 Seconds
(without alignments)
8344.847 Million cell updates/sec

Title: US-09-699-295-301
Perfect score: 1155
Sequence: 1 atggtggttgaggttgattc..........caagaaccagaaataaataa 1155

Scoring table: IDENTITY_NUC
Gapop 60.0 , Gapext 60.0

Searched: 11351937 seqs, 5372889281 residues

Word size : 0

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : EST:*
1: em_estfun:*
2: em_esthum:*
3: em_estin:*
4: em_estom:*
5: em_estpl:*
6: em_estba:*
7: em_estro:*
8: em_estov:*
9: em_htc:*
10: gb_est1:*
11: gb_est2:*
12: gb_htc:*
13: gb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_rod:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	141	12.2	865	11	BF676987	BF676987 602084215
	2	131	11.3	451	10	AI804733	AI804733 tq42b03.x
	3	129	11.2	531	13	AQ615477	AQ615477 HS_5144_B
	4	89	7.7	400	13	AQ124119	AQ124119 HS_3122_A
c	5	57	4.9	607	13	B48260	B48260 RPCI11-6K4.
c	6	52	4.5	380	11	BF329652	BF329652 RC6-BN027
	7	50	4.3	592	13	AQ372700	AQ372700 RPCI11-14
	8	47	4.1	187	10	BE069869	BE069869 CM1-BI039
	9	46	4.0	400	13	AQ057106	AQ057106 CIT-HSP-2
	10	45	3.9	894	11	BF675049	BF675049 602136643
c	11	41	3.5	503	13	B55862	B55862 CIT-HSP-200
	12	39	3.4	1011	13	AQ900910	AQ900910 HS_2055_B
	13	38	3.3	458	13	AQ247090	AQ247090 HS_2055_B
	14	37	3.2	694	11	BG720647	BG720647 [illegible]
c	15	32	2.8	443	13	AQ681878	AQ681878 HS_[illegible]_B
c	16	32	2.8	495	13	AQ469831	AQ469831 CITBI-E1
c	17	32	2.8	552	13	AQ469663	AQ469663 CITBI-E1
c	18	32	2.8	670	13	AQ344968	AQ344968 RPCI11-1[illegible]
c	19	32	2.8	697	13	AQ030113	AQ030113 RPCI11-[illegible]
c	20	32	2.8	875	13	AQ960245	AQ960245 HS_[illegible]_A
c	21	31	2.7	386	13	AQ134631	AQ134631 HS_[illegible]_B
c	22	31	2.7	388	13	AQ103336	AQ103336 HS_[illegible]_A
c	23	28	2.4	599	13	AQ531987	AQ531987 RPCI11-[illegible]
c	24	28	2.4	677	13	AQ385757	AQ385757 RPCI11-[illegible]
c	25	28	2.4	680	13	AQ779789	AQ779789 HS_[illegible]_B
	26	27	2.3	385	13	AQ063365	AQ063365 CIT-HSP-[illegible]
	27	27	2.3	471	13	AQ247094	AQ247094 HS_2055_B
c	28	26	2.3	543	13	AQ442715	AQ442715 HS_[illegible]_B
c	29	23	2.0	476	13	AQ392059	AQ392059 CITBI-E1
	30	21	1.8	399	11	BI265394	BI265394 NF[illegible]
	31	21	1.8	527	13	AQ622504	AQ622504 HS_[illegible]_A
	32	21	1.8	630	11	BI264745	BI264745 NIH_[illegible]
	33	21	1.8	675	11	BG454086	BG454086 NF[illegible]
	34	21	1.8	864	13	AQ746541	AQ746541 HS_[illegible]_A
	35	20	1.7	249	13	AZ472817	AZ472817 [illegible]
c	36	20	1.7	336	13	AQ697778	AQ697778 HS_[illegible]_A
	37	20	1.7	398	13	AQ211285	AQ211285 HS_[illegible]_A
c	38	20	1.7	403	11	BG042430	BG042430 [illegible]
	39	20	1.7	477	13	AQ596029	AQ596029 HS_2[illegible]_B
	40	20	1.7	488	13	AZ025907	AZ025907 RPCI[illegible]
c	41	20	1.7	502	13	AQ596426	AQ596426 HS_[illegible]_A
	42	20	1.7	595	13	AQ721770	AQ721770 HS_[illegible]_B
	43	20	1.7	646	13	AZ941453	AZ941453 2M0[illegible]
c	44	19	1.6	196	13	AZ784170	AZ784170 2M0[illegible]
	45	19	1.6	257	10	AI766969	AI766969 wi[illegible] x

ALIGNMENTS

RESULT 1
BF676987
LOCUS BF676987 865 bp mRNA EST [illegible]
DEFINITION 602084215F1 NIH_MGC_83 Homo sapiens cDNA clone IMAGE:4248746 5',
mRNA sequence.
ACCESSION BF676987
VERSION BF676987.1 GI:11950882
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 865)
AUTHORS NIH-MGC [illegible]
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: CLONETECH Laboratories, Inc.
cDNA Library Preparation: CLONETECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1067 row: m column: 03
High quality sequence stop: 642.
FEATURES Location/Qualifiers
source 1..865
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4248746"
/clone_lib="NIH_MGC_83"
/lab_host="DH10B (T1 phage resistant)"
/note="Organ: prostate; Vector: pDNR-LIB (Clontech)

Site_1: SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGGCCGAGGCGGCCGACATG-d(30)BN 3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.4 kb (range 0.5-4.0 kb). 14/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."
BASE COUNT 297 a 140 c 211 g 217 t
ORIGIN

Query Match 12.2%; Score 141; DB 11; Length 865;
Best Local Similarity 100.0%; Pred. No. 1.3e-62;
Matches 141; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy     915 actgctctcatacttgctgtatgttgtggatcagcaagtatagtcagcctctactgag 978
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     353 ACTGCTCTCATACTTGCTGTATGTTGTGGATCAGCAAGTATAGTCAGCCTCTACTGAG 412

Qy     979 caaaatattgatgtatcttctcaagatctatctggacagacggccagagagtatgctgtt 1038
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     413 CAAAATATTGATGTATCTTCTCAAGATCTATCTGGACAGACGGCCAGAGAGTATGCTGTT 472

Qy    1039 tctagtcatcatcatgtaatt 1059
           |||||||||||||||||||||
Db     473 TCTAGTCATCATCATGTAATT 493
```

RESULT 2
AI804733
LOCUS AI804733 451 bp mRNA EST 07-MAR-2000
DEFINITION [illegible] mRNA sequence.
ACCESSION AI804733
VERSION AI804733.1 GI:5470205
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 451)
AUTHORS NCI_CGAP http://www.ncbi.nlm.nih.gov/ncicgap
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 543 Std Error: 0.00
Seq primer: -40UP from Gibco.
FEATURES Location/Qualifiers
source 1..451
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2253677"
/clone_lib="NCI_CGAP_Pr28"
/sex="male"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: prostate; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Plasmid DNA from the normalized library NCI_CGAP_Pr22 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneIDs 985608-986759, 1101192-1101959, and 1217929-1220615). Subtraction by Bento Soares and M. Fatima Bonaldo."
BASE COUNT 128 a 79 c 91 g 153 t
ORIGIN

Query Match 12.1%; Score 140; DB 10; Length 451;
Best Local Similarity 100.0%; Pred. No. 4.3e-62;
Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy     916 aggactgctctcatacttgctgtatgttgtggatcagcaagtatagtcagcctctactt 975
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      47 AGGACTGCTCTCATACTTGCTGTATGTTGTGGATCAGCAAGTATAGTCAGCCTCTACTT 106

Qy     976 gaggcaaaatattgatgtatcttctcaagatctatctggacagacggccagagagtatgct 1035
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     107 GAGCAAAATATTGATGTATCTTCTCAAGATCTATCTGGACAGACGGCCAGAGAGTATGCT 166

Qy    1036 gtttctagtcatcatcatgt 1055
           ||||||||||||||||||||
Db     167 GTTTCTAGTCATCATCATGT 186
```

RESULT 3
AQ615477
LOCUS AQ615477 531 bp DNA GSS 15-JUN-1999
DEFINITION HS_5144_B1_G01_T7A RPCI-11 Human Male BAC Library Homo sapiens genomic clone Plate=720 Col=1 Row=N, DNA sequence.
ACCESSION AQ615477
VERSION AQ615477.1 GI:5076753
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 531)
AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,E., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.
TITLE Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
JOURNAL Proc. Natl. Acad. Sci. U. S. A. 93 (7), 3253-3744 (1999)
MEDLINE 99380589
COMMENT Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm) or from Research Genetics (info@resgen.com). BAC end Web Server: http://www.htsc.washington.edu
Plate: 720 row: N column: 1
Seq primer: T7
Class: BAC ends
High quality sequence stop: 531.
FEATURES Location/Qualifiers
source 1..531
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Plate=720 Col=1 Row=N"
/clone_lib="RPCI-11 Human Male BAC Library"
/sex="male"
/note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI; Male blood DNA was isolated from one randomly chosen donor and partially digested with a combination of EcoKI and EcoRI Methylase. Size selected DNA was cloned into the

```
                   pBACe3.6 vector at EcoRI sites"
BASE COUNT      167 a     92 c     92 g    169 t     11 others
ORIGIN

  Query Match             11.2%;  Score 12[illegible];  DB 13;  Length 531;
  Best Local Similarity   100.0%;  Pred. No. 2.5e-56;
  Matches  12[illegible];  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        682 catggcactgatccaaatattccagatgagtatggaaataccactctgcactacgctatc 741
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        117 CATGGCACTGATCCAAATATTCCAGATGAGTATGGAAATACCACTCTGCACTACGCTATC 176

Qy        742 tataatgaagataaattaatggccaaagcactgctcttatatggtgctgatatcgaatca 801
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        177 TATAATGAAGATAAATTAATGGCCAAAGCACTGCTCTTATATGGTGCTGATATCGAATCA 236

Qy        802 aaaaacaag 810
              |||||||||
Db        237 AAAAACAAG 245


RESULT  4
AQ124119
LOCUS       AQ124119       400 bp    DNA             GSS       22-SEP-1998
DEFINITION  HS_3122_A1_E07_MR CIT Approved Human Genomic Sperm Library D Homo
            sapiens genomic clone Plate=3122 Col=13 Row=E, DNA sequence.
ACCESSION   AQ124119
VERSION     AQ124119.1  GI:3501285
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 400)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., [illegible], [illegible], [illegible], [illegible], Adams,M.[illegible] and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
  MEDLINE   99380589
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Sequence Tagged Connector
            Plate: 3122  row: E  column: 13
            Class: BAC ends
            High quality sequence stop: 400.
FEATURES             Location/Qualifiers
     source          1..400
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="Plate=3122 Col=13 Row=E"
                     /clone_lib="CIT Approved Human Genomic Sperm Library D"
                     /sex="male"
                     /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                     E. Coli DH10B"
BASE COUNT      13[illegible] a     67 c     88 g    114 t
ORIGIN

  Query Match             7.7%;  Score 89;  DB 13;  Length 400;
  Best Local Similarity   100.0%;  Pred. No. 2.2e-35;
  Matches  89;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        6[illegible] [illegible]gatggcactgatc 694
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2[illegible] AGG[illegible]TACAATG[illegible]CAGGAAGATGAATG[illegible]G[illegible]TAATG[illegible]GCTGGAACATGGCACTGATC 296

Qy        695 caaatattccagatgagtatggaaatacc 723
              |||||||||||||||||||||||||||||
Db        297 CAAATATTCCAGATGAGTATGGAAATACC 325


RESULT  5
B48260/c
LOCUS       B48260         607 bp    DNA             GSS       0[illegible]-APR-199[illegible]
DEFINITION  RPCI11-6K4.TV RPCI-11 Homo sapiens genomic clone RPCI-11-6K4, DNA
            sequence.
ACCESSION   B48260
VERSION     B48260.1  GI:2600497
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 607)
  AUTHORS   Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., [illegible], [illegible]
            ,K., Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and Venter
            ,J.C.
  TITLE     Use of BAC End Sequences for Sequence Ready Map Building
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 208[illegible], USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genetics (info@resgen.com). BAC end search page
            http://www.[illegible]/bac_end_search/bac_end_search.html.
            Seq primer: T7
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1..607
                     /organism="Homo sapiens"
                     /db_xref="GDB:7502163"
                     /db_xref="taxon:9606"
                     /clone="RPCI-11-6K4"
                     /clone_lib="RPCI-11"
                     /sex="Male"
                     /cell_type="Lymphocytes"
                     /note="Vector: pBACe3.6; Site 1: EcoRI; Site 2: EcoRI;
                     RPCI11 Human Male BAC library"
BASE COUNT      186 a    131 c    109 g    181 t
ORIGIN

  Query Match             4.9%;  Score 57;  DB 13;  Length 607;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-18;
  Matches  57;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        575 [illegible] 631
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        148 [illegible] 9[illegible]


RESULT  6
BF329652/c
LOCUS       BF329652       380 bp    mRNA            EST       2[illegible]-NOV-[illegible]
DEFINITION  RC5-BN0276-160600-011-F12 BN0276 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BF329652
VERSION     BF329652.1  GI:1[illegible]
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 380)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar  01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=BN0276-RPCI-BN0276-
            160599-011-F12&t2=2000-06-16&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 99
            High quality sequence stop: 379.
FEATURES             Location/Qualifiers
     source          1. .380
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="BN0276"
                     /dev_stage="Adult"
                     /note "Organ: breast_normal; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
BASE COUNT       80 a    106 c    109 g     85 t
ORIGIN


  Query Match             4.5%;  Score 52;  DB 11;  Length 380;
  Best Local Similarity   100.0%;  Pred. No. 5.4e-16;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     296 gcaagtggtgctgccactgcttcccctgctgcaggggaagcggcaaagagcaa 347
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     232 GCAAGTGGTGCTGCCACTGCTTCCCCTGCTGCAGGGGAAGCGGCAAAGAGCAA 181


RESULT   7
AQ372700
LOCUS       AQ372700       592 bp    DNA             GSS       20-MAY-1999
DEFINITION  RPCI11-147I2.TJ RPCI-11 Homo sapiens genomic clone RPCI 11 147I2,
            DNA sequence.
ACCESSION   AQ372700
VERSION     AQ372700.1  GI:4443723
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 592)
  AUTHORS   Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and Venter
            ,J.C.
  TITLE     Use of BAC End Sequences from Library RPCI-11 for Sequence Ready
            Map Building
  JOURNAL   Unpublished (1997)
COMMENT     Other_GSSs: RPCI11-147I2.1V
            Contact: Shaying Zhao, William Nierman, Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: hbe@tigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genetics (info@resgen.com). BAC end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_e_search.html
            Seq primer: SP6
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .592
                     /organism="Homo sapiens"
                     /db_xref="GDB:7556257"
                     /db_xref="taxon:9606"
                     /clone="RPCI-11-147I2"
                     /clone_lib="RPCI-11"
                     /sex="Male"
                     /cell_type "lymphocytes"
                     /note "Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     RPCI11 Human Male BAC Library"
BASE COUNT      199 a     87 c     99 g    207 t
ORIGIN


  Query Match             4.3%;  Score 50;  DB 13;  Length 592;
  Best Local Similarity   100.0%;  Pred. No. 6.1e-15;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     745 aatgaagataaattaatggccaaaagcactgctcttatatggtgctgatat 794
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     153 AATGAAGATAAATTAATGGCCAAAGCACTGCTCTTATATGGTGCTGATAT 202


RESULT   8
BE069869
LOCUS       BE069869       187 bp    mRNA            EST       04-JUN-2000
DEFINITION  CM1 BT0397 201299-073 a11 BT0397 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BE069869
VERSION     BE069869.1  GI:8414519
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 187)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=CM1-BT0397-201
            299-073-a11&t3=1999-12-20&t4=1)
            Seq primer: puc 18 forward
```

```
                 High quality sequence start: 82
                 High quality sequence stop: 141.
FEATURES                  Location/Qualifiers
     source               1..187
                          /organism "Homo sapiens"
                          /db_xref "taxon:9606"
                          /clone_lib "BI0697"
                          /dev_stage "Adult"
                          /note "Organ: breast; Vector: puc18; Site_1: SmaI; Site_2:
                          SmaI; A mini library was made by cloning products derived
                          from ORESTES PCR (U.S. Letters Patent application No. 196
                          ,716 - Ludwig Institute for Cancer Research) profiles
                          into the pUC 18 vector. Reverse transcription of tissue
                          mRNA and cDNA amplification were performed under low
                          stringency conditions."
BASE COUNT        66 a     45 c     34 g     42 t
ORIGIN

  Query Match             4.1%;  Score 47;  DB 10;  Length 187;
  Best Local Similarity   100.0%;  Pred. No. 2.2e-13;
  Matches   47;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        790 gatatgaatcaaaaacaagcatggcctcacaccactgttacttgg 836
              [illegible]
Db         91 GATATGAATCAAAAACAAGCATGGCCTCACACCACTGTTACTTGG 137


RESULT 9
AQ057106
LOCUS       AQ057106        400 bp    DNA             GSS       30-JUL-1998
DEFINITION  CIT-HSP-2335J19.TF CIT-HSP Homo sapiens genomic clone 2335J19, DNA
            sequence.
ACCESSION   AQ057106
VERSION     AQ057106.1  GI:[illegible]
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 400)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,C., Fisher,K.,
            Golden,K., Berry,K., Crandell,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence Ready Map
            Building (1998)
  JOURNAL   Unpublished (1998)
COMMENT     Other GSSs: CIT-HSP-2335J19.TR
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page
            http://www.[illegible]/bac_end_search.html.
            Seq primer: M13-21
            Class: BAC ends.
FEATURES                  Location/Qualifiers
     source               1..400
                          /organism "Homo sapiens"
                          /db_xref "taxon:9606"
                          /clone "2335J19"
                          /clone_lib "CIT-HSP"
                          /sex "Male"
                          /cell_type "Sperm"
                          /note "Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                          HindIII"
BASE COUNT       124 a     61 c     83 g    132 t
ORIGIN

  Query Match             4.0%;  Score 46;  DB 13;  Length 400;
  Best Local Similarity   100.0%;  Pred. No. 7.5e-13;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1056 aatttgccagttactttctgactacaaagaaaaacagatgctaaaa 1101
              ||||||||||||||||||||||||||||||||||||||||||||||
Db        219 AATTTGCCAGTTACTTTCTGACTACAAAGAAAAACAGATGCTAAAA 264


RESULT 10
BF675049
LOCUS       BF675049        894 bp    mRNA            EST       21-DEC-2000
DEFINITION  602136643F1 NIH_MGC_83 Homo sapiens cDNA clone IMAGE:4273139 5',
            mRNA sequence.
ACCESSION   BF675049
VERSION     BF675049.1  GI:11948944
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 894)
  AUTHORS   NIH-MGC http://[illegible]
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: CLONETECH Laboratories, Inc.
             cDNA Library Preparation: CLONETECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1088  row: e  column: 12
            High quality sequence start: 6
            High quality sequence stop: 576.
FEATURES                  Location/Qualifiers
     source               1..894
                          /organism "Homo sapiens"
                          /db_xref "taxon:9606"
                          /clone "IMAGE:4273139"
                          /clone_lib "NIH_MGC_83"
                          /lab_host "DH10B (T1 phage resistant)"
                          /note "Organ: prostate; Vector: pDNR-LIB (Clontech)
                          Site_1: SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc
                          ); 5' and 3' adaptors were used in cloning as follows: 5'
                          adaptor sequence: 5' CACGGCCATTATGGC 3' and 3' adaptor
                          sequence: 5' ATTCTAGAGGCCGAGGCGGCCGACATG d(T)30VN 3'
                          (where V = A, C, or G and N = A, C, G, or T). Average
                          insert size 1.4 kb (range 0.5-4.0 kb). 14/15 colonies
                          contained inserts by PCR. This library was enriched for
                          full length clones and was constructed by Clontech
                          Laboratories (Palo Alto, CA)."
BASE COUNT       239 a    186 c    208 g    261 t
ORIGIN

  Query Match             3.9%;  Score 45;  DB 11;  Length 894;
  Best Local Similarity   100.0%;  Pred. No. 2.6e-12;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        651 ggaagatgaatgtgcgttaatgttgctggaacatggcactgatcc [illegible]
              |||||||||||||||||||||||||||||||||||||||||||||
Db         46 GGAAGATGAATGTGCGTTAATGTTGCTGGAACATGGCACTGATCC [illegible]


RESULT 11
B55862/c
LOCUS       B55862          503 bp    DNA             GSS       20-JUN-1998
DEFINITION  CIT-HSP-2005F16.TF CIT-HSP Homo sapiens genomic clone [illegible], DNA
            sequence.
```

```
ACCESSION   B55862
VERSION     B55862.1  GI:2610196
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo
REFERENCE   1  (bases 1 to 503)
  AUTHORS   Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K., Golden
            ,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M.
            and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence Read, Map-
            Building
  JOURNAL   Unpublished (1997)
COMMENT     Other_GSSs: CIT-HSP-2005F16 TR
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: M13-21
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1..503
                     /organism="Homo sapiens"
                     /db_xref="UI: 20242"
                     /db_xref="taxon:9606"
                     /clone="2005F16"
                     /clone_lib="CIT-HSP"
                     /sex="Male"
                     /cell_type="Sperm"
                     /note "Vector: pBeloBAC11, Site_1: HindIII, Site_2:
                     HindIII"
BASE COUNT      152 a    116 c     87 g    148 t
ORIGIN


  Query Match              3.5%;  Score 40;  DB 13;  Length 503;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-09;
  Matches   40;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    1087 aaacagatgctaaaaatctcttctgaaaacagcaatccag 1126
           ||||||||||||||||||||||||||||||||||||||||||
Db     502 AAACAGATGCTAAAAATCTCTTCTGAAAACAGCAATCCAG 463


RESULT  12
AQ900910
LOCUS       AQ900910      1011 bp     DNA              GSS       10-NOV-1999
DEFINITION  HS_2055_B1_B05_F7C CIT Approved Human Genomic Sperm Library D Homo
            sapiens genomic clone Plate=2055 Col=9 Row=D, DNA sequence.
ACCESSION   AQ900910
VERSION     AQ900910.1  GI:6357012
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1011)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
  MEDLINE   99380589
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Clones may be purchased from Research Genetics (info@resgen.com).
            BAC end Web Server: http://www.htsc.washington.edu
            Plate: 2055  row: D  column: 9
            Seq primer: T7
            Class: BAC ends
            High quality sequence stop: 1011.
FEATURES             Location/Qualifiers
     source          1..1011
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="Plate=2055 Col=9 Row=D"
                     /clone_lib="CIT Approved Human Genomic Sperm Library D"
                     /sex="male"
                     /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                     E-Coli DH10B"
BASE COUNT      216 a    214 c    143 g    438 t
ORIGIN


  Query Match              3.4%;  Score 39;  DB 13;  Length 1011;
  Best Local Similarity   100.0%;  Pred. No. 3.6e-09;
  Matches   39;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    1054 gtaatttgccaqttacttttctgactacaaaqaaaaaacaq 1092
           |||||||||||||||||||||||||||||||||||||||||||
Db     237 GTAATTTGCCAGTTACTTTCTGACTACAAAGAAAAACAG 245


RESULT  13
AQ247090
LOCUS       AQ247090      458 bp     DNA              GSS       06-OCT-1998
DEFINITION  HS_2055_B1_B04_F7 CIT Approved Human Genomic Sperm Library D Homo
            sapiens genomic clone Plate=2055 Col=7 Row=D, DNA sequence.
ACCESSION   AQ247090
VERSION     AQ247090.1  GI:3697272
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 458)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
  MEDLINE   99380589
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Sequence Tagged Connector
            Plate: 2055  row: D  column: 7
            Class: BAC ends
            High quality sequence stop: 458.
FEATURES             Location/Qualifiers
     source          1..458
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="Plate=2055 Col 7 Row=D"
                     /clone_lib "CIT Approved Human Genomic Sperm Library D"
                     /sex="male"
                     /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                     E-Coli DH10B"
```

```
BASE COUNT      110 a    122 c     68 g    158 t
ORIGIN

  Query Match             3.3%;  Score 38;  DB 13;  Length 458;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-08;
  Matches   38;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1088 aacagatgctaaaaatctcttctgaaacagcaatcca 1125
              ||||||||||||||||||||||||||||||||||||||
Db        198 AACAGATGCTAAAAATCTCTTCTGAAACAGCAATCCA 235


RESULT 14
BG720647
LOCUS       BG720647        694 bp    mRNA            EST       08-MAY-2001
DEFINITION  602592528F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4824630 5',
            mRNA sequence.
ACCESSION   BG720647
VERSION     BG720647.1  GI:13999834
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 694)
  AUTHORS   [illegible]
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
            cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10736  row: h  column: 07
            High quality sequence stop: 694.
FEATURES             Location/Qualifiers
     source          1..694
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4824630"
                     /clone_lib="NIH_MGC_97"
                     /lab_host="DH10B"
                     /note="Organ: testis; Vector: pBluescriptR (modified
                     pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
                     ); oligo dT primed using primer 5' TTTTTTTTTTTTTTTTTTVN-3',
                     size selected for  average insert size 2.2 kb and
                     normalized to ROT 5. This is a primary library enriched
                     for full length clones and constructed using the
                     Cap trapper method (Carninci, in preparation).  Library
                     constructed by M. Brownstein (NIMH/NHGRI, National
                     Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT      108 a    173 c    228 g    185 t
ORIGIN

  Query Match             3.2%;  Score 37;  DB 11;  Length 694;
  Best Local Similarity   100.0%;  Pred. No. 4e-08;
  Matches   37;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        159 tatgaagacactcaggagcaagatgggcaagtggtgc 195
              |||||||||||||||||||||||||||||||||||||
Db        645 TATGAAGACACTCAGGAGCAAGATGGGCAAGTGGTGC 681


RESULT 15
AQ681878/c
LOCUS       AQ681878        443 bp    DNA             GSS       28-JUN-1999
DEFINITION  HS_5500_B2_B06_T7A RPCI-11 Human Male BAC Library Homo sapiens
            genomic clone Plate=1076 Col=12 Row=D, DNA sequence.
ACCESSION   AQ681878
VERSION     AQ681878.1  GI:5257861
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 443)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
  MEDLINE   99380589
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
            or from Research Genetics (info@resgen.com). BAC end Web Server:
            http://www.htsc.washington.edu
            Plate: 1076  row: D  column: 12
            Seq primer: T7
            Class: BAC ends
            High quality sequence stop: 443.
FEATURES             Location/Qualifiers
     source          1..443
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="Plate=1076 Col=12 Row=D"
                     /clone_lib="RPCI-11 Human Male BAC Library"
                     /sex="male"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     Male blood DNA was isolated from one randomly chosen donor
                     and partially digested with a combination of EcoRI and
                     EcoRI Methylase. Size selected DNA was cloned into the
                     pBACe3.6 vector at EcoRI sites"
BASE COUNT      158 a     62 c     56 g    164 t      3 others
ORIGIN

  Query Match             2.8%;  Score 32;  DB 13;  Length 443;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-05;
  Matches   32;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1124 cagaaaatgtctcaagaaccagaaataaataa 1155
              ||||||||||||||||||||||||||||||||
Db        217 CAGAAAATGTCTCAAGAACCAGAAATAAATAA 186


Search completed: March 9, 2002, 03:42:43
Job time: 4971 sec
```

```
OM nucleic - nucleic search, using sw model

Run on:          March 9, 2002, 02:21:32 ; Search time 1718.77 Seconds
                                           (without alignments)
                                           11085.976 Million cell updates/sec

Title:           US-09-699-295-301
Perfect score:   1155
Sequence:        1 atggtggttgaggttgattc..........caagaaccagaaataaataa 1155

Scoring table:   OLIGO NUC
                 Gapop 60.0 , Gapext 60.0

Searched:        1472140 seqs, 8248589755 residues

Word size :      0

Total number of hits satisfying chosen parameters:       2944280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :       GenEmbl:*
                1:  gb_ba:*
                2:  gb_htg:*
                3:  gb_in:*
                4:  gb_om:*
                5:  gb_ov:*
                6:  gb_pat:*
                7:  gb_ph:*
                8:  gb_pl:*
                9:  gb_pr:*
                10:  gb_ro:*
                11:  gb_sts:*
                12:  gb_sy:*
                13:  gb_un:*
                14:  gb_vi:*
                15:  em_ba:*
                16:  em_fun:*
                17:  em_hum:*
                18:  em_in:*
                19:  em_om:*
                20:  em_or:*
                21:  em_ov:*
                22:  em_pat:*
                23:  em_ph:*
                24:  em_pl:*
                25:  em_ro:*
                26:  em_sts:*
                27:  em_sy:*
                28:  em_un:*
                29:  em_vi:*
                30:  em_htgo_hum:*
                31:  em_htgo_inv:*
                32:  em_htgo_rod:*
                33:  em_htg_hum:*
                34:  em_htg_inv:*
                35:  em_htg_rod:*
                36:  em_htg_other:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	Query Match	Length	DB	ID	Description
	1	1155	100.0	1155	6	AX106592	AX106592 Sequence
	2	1155	100.0	1155	6	AX140883	AX140883 Sequence
	3	1128	97.7	2000	6	AX106593	AX106593 Sequence
	4	1128	97.7	2000	6	AX140884	AX140884 Sequence
	5	1128	97.7	2040	6	AX106594	AX106594 Sequence
	6	1128	97.7	2040	6	AX140885	AX140885 Sequence
	7	491	42.5	1512	6	AX106587	AX106587 Sequence
	8	491	42.5	1512	6	AX140878	AX140878 Sequence
c	9	453	39.2	199668	2	AC018804	AC018804 Homo sapi
	10	252	21.8	1853	6	AX106588	AX106588 Sequence
	11	252	21.8	1853	6	AX140879	AX140879 Sequence
c	12	137	11.9	115727	2	AC013269	AC013269 Homo sapi
	13	120	10.4	879	6	AX141041	AX141041 Sequence
	14	120	10.4	1059	6	AX106591	AX106591 Sequence
	15	120	10.4	1059	6	AX140882	AX140882 Sequence
c	16	120	10.4	1851	6	AR148112	AR148112 Sequence
c	17	120	10.4	1851	6	AX106585	AX106585 Sequence
c	18	120	10.4	1851	6	AX140876	AX140876 Sequence
	19	120	10.4	1852	6	AX141040	AX141040 Sequence
	20	120	10.4	1855	6	AX106590	AX106590 Sequence
	21	120	10.4	1855	6	AX140881	AX140881 Sequence
	22	120	10.4	2184	6	AX106589	AX106589 Sequence
	23	120	10.4	2184	6	AX140880	AX140880 Sequence
c	24	118	10.2	668	6	AX106586	AX106586 Sequence
c	25	118	10.2	668	6	AX140877	AX140877 Sequence
c	26	109	9.4	454	6	AR148032	AR148032 Sequence
	27	94	8.1	167293	2	AC022629	AC022629 Homo sapi
c	28	94	8.1	177312	2	AC036207	AC036207 Homo sapi
	29	87	7.5	83276	2	AC026631	AC026631 Homo sapi
	30	86	7.4	80955	2	AC026984	AC026984 Homo sapi
	31	86	7.4	104785	9	AP001465	AP001465 Homo sapi
	32	86	7.4	153077	9	AC020679	AC020679 Homo sapi
	33	86	7.4	172550	2	AC060814	AC060814 Homo sapi
	34	86	7.4	184109	2	AC023191	AC023191 Homo sapi
c	35	86	7.4	217377	2	AC037471	AC037471 Homo sapi
	36	86	7.4	340000	9	HS21C003	AL163203 Homo sapi
c	37	80	6.9	154077	9	CNS07EFD	AL512024 Human chr
	38	71	6.1	115727	2	AC013269	AC013269 Homo sapi
	39	71	6.1	121963	2	AC018865	AC018865 Homo sapi
c	40	61	5.3	37981	9	AP000529	AP000529 Homo sapi
	41	50	4.3	72871	2	AC091550	AC091550 Homo sapi
c	42	46	4.0	184109	2	AC023191	AC023191 Homo sapi
c	43	45	3.9	69992	2	AC067905	AC067905 Homo sapi
c	44	43	3.7	173957	2	AC091671	AC091671 Papio cyn
c	45	40	3.5	61407	2	AC012414	AC012414 Homo sapi

ALIGNMENTS

```
RESULT 1
AX106592
LOCUS       AX106592      1155 bp    DNA             PAT       30-APR-2001
DEFINITION  Sequence 373 from Patent WO0125272.
ACCESSION   AX106592
VERSION     AX106592.1  GI:13922263
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1155)
  AUTHORS   Xu,J., Skeiky,Y.A., Reed,S.G. and Cheever,M.A.
  TITLE     Compositions and methods for therapy and diagnosis of prostate
            cancer
  JOURNAL   Patent: WO 0125272-A 373 12-APR-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..1155
                     /organism="Homo sapiens"
```

```
                     /db_xref="taxon:9606"
BASE COUNT      346 a    253 c    297 g    259 t
ORIGIN


  Query Match             100.0%;  Score 1155;  DB 6;  Length 1155;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1155;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 atggtggttgaggttgattccatgccggctgcctcttctgtgaagaagccatttggtctc 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ATGGTGGTTGAGGTTGATTCCATGCCGGCTGCCTCTTCTGTGAAGAAGCCATTTGGTCTC 60

Qy      61 aggaggcaaggatgggcaagtggtgctgccgttgcttccccctgctgcagggagaggcggcaag 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 AGGAGGCAAGATGGGCAAGTGGTGCTGCCGTTGCTTCCCCTGCTGCAGGGAGAGCGGCAAG 120

Qy     121 agcaacgtgggcacttctggagaccacgacgactctgctatgaagacactcaggaacaagat 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 AGCAACGTGGGCACTTCTGGAGACCACGACGACTCTGCTATGAAGACACTCAGGAACAAGAT 180

Qy     181 atgggcaagtggtgccgccactgcttccccctgctgcagggggagtggcaagagcaacgtg 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 ATGGGCAAGTGGTGCCGCCACTGCTTCCCCTGCTGCAGGGGGAGTGGCAAGAGCAACGTG 240

Qy     241 ggcgcttctggagaccacgacgactctgctatgaagacactcaggaacaagatgggcaag 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 GGCGCTTCTGGAGACCACGACGACTCTGCTATGAAGACACTCAGGAACAAGATGGGCAAG 300

Qy     301 tggtgctgccgctgcttccccctgctgcagggggagcggcaagagcaacgtgggcgcttgg 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 TGGTGCTGCCGCTGCTTCCCCTGCTGCAGGGGGAGCGGCAAGAGCAACGTGGGCGCTTGG 360

Qy     361 ggaggactacgatgacgagtgccttcatggaggccaggtaccacgtccgtggaggaagatctg 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 GGAGGACTACGATGACGAGTGCCTTCATGGAGGCCAGGTACCACGTCCGTGGAGGAAGATCTG 420

Qy     421 gacaagctccacagaggctgcctggtgggggtaaaggtccccagaaggcatctcatcggccatg 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GACAAGCTCCACAGAGGCTGCCTGGTGGGGGTAAAGGTCCCCAGAAGGCATCTCATCGGCATG 480

Qy     481 ctcagggacactgacgtgaacaagaaggacaagcaaaagaggactgctctacatctggcc 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 CTCAGGGACACTGACGTGAACAAGAAGGACAAGCAAAAGAGGACTGCTCTACATCTGGCC 540

Qy     541 tctgccaatgggaattcagaagtagtaaaactcctgctggacagagcgatgtcaacttaat 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 TCTGCCAATGGGAATTCAGAAGTAGTAAAACTCCTGCTGGACAGAGCGATGTCAACTTAAT 600

Qy     601 gtccttgacaacaaaaagaggacagctctgataaaggccgtacaatgccaggaagatgaa 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 GTCCTTGACAACAAAAAGAGGACAGCTCTGATAAAGGCCGTACAATGCCAGGAAGATGAA 660

Qy     661 tgtgcgttaatgttgctggaacatggcactgatccaaatattccagatgagtatggaaat 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 TGTGCGTTAATGTTGCTGGAACATGGCACTGATCCAAATATTCCAGATGAGTATGGAAAT 720

Qy     721 accactctgcactacgctatctataatgaagataaattaatggccaaagcactgctctta 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 ACCACTCTGCACTACGCTATCTATAATGAAGATAAATTAATGGCCAAAGCACTGCTCTTA 780

Qy     781 tatggtgctgatatcgaatcaaaaaacaagcatggcctcacaccactgttacttggtgta 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 TATGGTGCTGATATCGAATCAAAAAACAAGCATGGCCTCACACCACTGTTACTTGGTGTA 840

Qy     841 catgagcaaaaacagcaagtcgtgaaatttttaatcaagaaaaaagcgaatttaaatgca 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 CATGAGCAAAAACAGCAAGTCGTGAAATTTTTAATCAAGAAAAAAGCGAATTTAAATGCA 900

Qy     901 ctggatagatatggaaggactgctctcatacttgctgtatgttgtggatcagcaagtata 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 CTGGATAGATATGGAAGGACTGCTCTCATACTTGCTGTATGTTGTGGATCAGCAAGTATA 960

Qy     961 gtcagccttctacttgaagcaaaaatattgatgtatcttctcaagatctatctggacagacg 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 GTCAGCCTTCTACTTGAGCAAAAATATTGATGTATCTTCTCAAGATCTATCTGGACAGACG 1020

Qy    1021 gccagaggagtatgctgtttctagtcatcatcatgtaatttgccagttactttctgactac 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 GCCAGAGGAGTATGCTGTTTCTAGTCATCATCATGTAATTTGCCAGTTACTTTCTGACTAC 1080

Qy    1081 aaagaaaaacagatgctaaaaatctcttctgaaaacagcaatccagaaaatgtctcaaga 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 AAAGAAAAACAGATGCTAAAAATCTCTTCTGAAAACAGCAATCCAGAAAATGTCTCAAGA 1140

Qy    1141 accagaaaataaataa 1155
           ||||||||||||||||
Db    1141 ACCAGAAATAAATAA 1155


RESULT   2
AX140883
LOCUS       AX140883     1155 bp    DNA             PAT       31-MAY-2001
DEFINITION  Sequence 373 from Patent WO0134802.
ACCESSION   AX140883
VERSION     AX140883.1  GI:14280986
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1155)
  AUTHORS   Xu,J., Dillon,D.C., Mitcham,J.L., Ballinger,S.J., Jiang,Y.,
            Reed,S.G., Kalos,M.D., Retter,M.W., Stolk,J.A., Day,C.H.,
            Skeiky,Y.A. and Wang,A.
  TITLE     Compositions and methods for the therapy and diagnosis of prostate
            cancer
  JOURNAL   Patent: WO 0134802 A 373 17-MAY-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..1155
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      346 a    253 c    297 g    259 t
ORIGIN


  Query Match             100.0%;  Score 1155;  DB 6;  Length 1155;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1155;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 atggtggttgaggttgattccatgccggctgcctcttctgtgaagaagccatttggtctc 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ATGGTGGTTGAGGTTGATTCCATGCCGGCTGCCTCTTCTGTGAAGAAGCCATTTGGTCTC 60

Qy      61 aggaggcaaggatgggcaagtggtgctgccgttgcttccccctgctgcagggagaggcggcaag 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 AGGAGGCAAGATGGGCAAGTGGTGCTGCCGTTGCTTCCCCTGCTGCAGGGAGAGCGGCAAG 120

Qy     121 agcaacgtgggcacttctggagaccacgacgactctgctatgaagacactcaggaacaagat 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 AGCAACGTGGGCACTTCTGGAGACCACGACGACTCTGCTATGAAGACACTCAGGAACAAGAT 180

Qy     181 atgggcaagtggtgccgccactgcttccccctgctgcagggggagtggcaagagcaacgtg 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 ATGGGCAAGTGGTGCCGCCACTGCTTCCCCTGCTGCAGGGGGAGTGGCAAGAGCAACGTG 240

Qy     241 ggcgcttctggagaccacgacgactctgctatgaagacactcaggaacaagatgggcaag 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 GGCGCTTCTGGAGACCACGACGACTCTGCTATGAAGACACTCAGGAACAAGATGGGCAAG 300

Qy     301 tggtgctgccgctgcttccccctgctgcagggggagcggcaagagcaacgtgggcgcttgg 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 TGGTGCTGCCGCTGCTTCCCCTGCTGCAGGGGGAGCGGCAAGAGCAACGTGGGCGCTTGG 360

Qy     361 ggaggactacgatgacgagtgccttcatggaggccaggtaccacgtccgtggaggaagatctg 420
```

```
Db    361 GGAGACTACGATGACAGTGCCTTCATGGAGCCCAGGTACCACGTCCGTGGAGAAGATCTG 420

Qy    421 gacaagctccacagagctgcctggtggggtaaagtccccagaaaggatctcatcgtcatg 480
Db    421 GACAAGCTCCACAGAGCTGCCTGGTGGGGTAAAGTCCCCAGAAAGGATCTCATCGTCATG 480

Qy    481 ctcagggacactgacgtgaacaagaaggacaagcaaaagaggactgctctacatctggcc 540
Db    481 CTCAGGGACACTGACGTGAACAAGAAGGACAAGCAAAAGAGGACTGCTCTACATCTGGCC 540

Qy    541 tctgccaatgggaattcagaagtagtaaaactcctgctggacagacgatgtcaacttaat 600
Db    541 TCTGCCAATGGGAATTCAGAAGTAGTAAAACTCCTGCTGGACAGACGATGTCAACTTAAT 600

Qy    601 gtccttgacaacaaaaagaggacagctctgataaaggccgtacaatgccaggaagatgaa 660
Db    601 GTCCTTGACAACAAAAAGAGGACAGCTCTGATAAAGGCCGTACAATGCCAGGAAGATGAA 660

Qy    661 tgtgcgttaatgttgctggaacatggcactgatccaaatattccagatgagtatggaaat 720
Db    661 TGTGCGTTAATGTTGCTGGAACATGGCACTGATCCAAATATTCCAGATGAGTATGGAAAT 720

Qy    721 accactctgcactacgctatctataatgaagataaattaatggccaaagcactgctctta 780
Db    721 ACCACTCTGCACTACGCTATCTATAATGAAGATAAATTAATGGCCAAAGCACTGCTCTTA 780

Qy    781 tatggtgctgatatcgaatcaaaaaacaagcatggcctcacaccactgttacttggtgta 840
Db    781 TATGGTGCTGATATCGAATCAAAAAACAAGCATGGCCTCACACCACTGTTACTTGGTGTA 840

Qy    841 catgagcaaaaacagcaagtcgtgaaatttttaatcaagaaaaaagcgaatttaaatgca 900
Db    841 CATGAGCAAAAACAGCAAGTCGTGAAATTTTTAATCAAGAAAAAAGCGAATTTAAATGCA 900

Qy    901 ctggatagatatggaaggactgctctcatacttgctgtatgttgtggatcagcaagtata 960
Db    901 CTGGATAGATATGGAAGGACTGCTCTCATACTTGCTGTATGTTGTGGATCAGCAAGTATA 960

Qy    961 gtcagccttctacttgagcaaaatattgatgtatcttctcaagatctatctggacagacg 1020
Db    961 GTCAGCCTTCTACTTGAGCAAAATATTGATGTATCTTCTCAAGATCTATCTGGACAGACG 1020

Qy   1021 gccagagagtatgctgtttctagtcatcatcatgtaatttgccagttactttctgactac 1080
Db   1021 GCCAGAGAGTATGCTGTTTCTAGTCATCATCATGTAATTTGCCAGTTACTTTCTGACTAC 1080

Qy   1081 aaagaaaaacagatgctaaaaatctcttctgaaaacagcaatccagaaaatgtctcaaga 1140
Db   1081 AAAGAAAAACAGATGCTAAAAATCTCTTCTGAAAACAGCAATCCAGAAAATGTCTCAAGA 1140

Qy   1141 accagaaataaataa 1155
Db   1141 ACCAGAAATAAATAA 1155
```

```
RESULT 3
AX106593
LOCUS       AX106593       2000 bp    DNA             PAT       30-APR-2001
DEFINITION  Sequence 374 from Patent WO0125272.
ACCESSION   AX106593
VERSION     AX106593.1  GI:13922264
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2000)
  AUTHORS   Xu,J., Skeiky,Y.A., Reed,S.G. and Cheever,M.A.
  TITLE     Compositions and methods for therapy and diagnosis of prostate
            cancer
  JOURNAL   Patent: WO 0125272-A 374 12-APR-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..2000
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      698 a    388 c    489 g    425 t
ORIGIN

  Query Match             97.7%;  Score 1128;  DB 6;  Length 2000;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1128;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 atggtggttgaggttgattccatgccggctgcctcttctgtgaagaagccatttggtctc 60
Db      1 ATGGTGGTTGAGGTTGATTCCATGCCGGCTGCCTCTTCTGTGAAGAAGCCATTTGGTCTC 60

Qy     61 aggagcaagatgggcaagtggtgctgccgttgcttcccctgctgcagggggagcggcaag 120
Db     61 AGGAGCAAGATGGGCAAGTGGTGCTGCCGTTGCTTCCCCTGCTGCAGGGGGAGCGGCAAG 120

Qy    121 agcaacgtgggcacttctggagaccacgacgactctgctatgaagacactcaggagcaag 180
Db    121 AGCAACGTGGGCACTTCTGGAGACCACGACGACTCTGCTATGAAGACACTCAGGAGCAAG 180

Qy    181 atgggcaagtggtgccgccactgcttcccctgctgcagggggagtggcaagagcaacgtg 240
Db    181 ATGGGCAAGTGGTGCCGCCACTGCTTCCCCTGCTGCAGGGGGAGTGGCAAGAGCAACGTG 240

Qy    241 ggcgcttctggagaccacgacgactctgctatgaagacactcaggagcaagatgggcaag 300
Db    241 GGCGCTTCTGGAGACCACGACGACTCTGCTATGAAGACACTCAGGAGCAAGATGGGCAAG 300

Qy    301 tggtgctgccactgcttcccctgctgcagggggagcggcaagagcaaggtgggcgcttgg 360
Db    301 TGGTGCTGCCACTGCTTCCCCTGCTGCAGGGGGAGCGGCAAGAGCAAGGTGGGCGCTTGG 360

Qy    361 ggagactacgatgacagtgccttcatggagcccaggtaccacgtccgtggagaagatctg 420
Db    361 GGAGACTACGATGACAGTGCCTTCATGGAGCCCAGGTACCACGTCCGTGGAGAAGATCTG 420

Qy    421 gacaagctccacagagctgcctggtggggtaaagtccccagaaaggatctcatcgtcatg 480
Db    421 GACAAGCTCCACAGAGCTGCCTGGTGGGGTAAAGTCCCCAGAAAGGATCTCATCGTCATG 480

Qy    481 ctcagggacactgacgtgaacaagaaggacaagcaaaagaggactgctctacatctggcc 540
Db    481 CTCAGGGACACTGACGTGAACAAGAAGGACAAGCAAAAGAGGACTGCTCTACATCTGGCC 540

Qy    541 tctgccaatgggaattcagaagtagtaaaactcctgctggacagacgatgtcaacttaat 600
Db    541 TCTGCCAATGGGAATTCAGAAGTAGTAAAACTCCTGCTGGACAGACGATGTCAACTTAAT 600

Qy    601 gtccttgacaacaaaaagaggacagctctgataaaggccgtacaatgccaggaagatgaa 660
Db    601 GTCCTTGACAACAAAAAGAGGACAGCTCTGATAAAGGCCGTACAATGCCAGGAAGATGAA 660

Qy    661 tgtgcgttaatgttgctggaacatggcactgatccaaatattccagatgagtatggaaat 720
Db    661 TGTGCGTTAATGTTGCTGGAACATGGCACTGATCCAAATATTCCAGATGAGTATGGAAAT 720

Qy    721 accactctgcactacgctatctataatgaagataaattaatggccaaagcactgctctta 780
Db    721 ACCACTCTGCACTACGCTATCTATAATGAAGATAAATTAATGGCCAAAGCACTGCTCTTA 780

Qy    781 tatggtgctgatatcgaatcaaaaaacaagcatggcctcacaccactgttacttggtgta 840
Db    781 TATGGTGCTGATATCGAATCAAAAAACAAGCATGGCCTCACACCACTGTTACTTGGTGTA 840

Qy    841 catgagcaaaaacagcaagtcgtgaaatttttaatcaagaaaaaagcgaatttaaatgca 900
Db    841 CATGAGCAAAAACAGCAAGTCGTGAAATTTTTAATCAAGAAAAAAGCGAATTTAAATGCA 900

Qy    901 ctggatagatatggaaggactgctctcatacttgctgtatgttgtggatcagcaagtata 960
```

```
Db    901 CTGGATAGATATGGAAGGACTGCTCTCATACTTGCTGIATGTTGTGGATCAGCAAGTATA 960

Qy    961 gtcagccttctacttgagcaaaatattgatgtatcttctcaagatctatctggacagacg 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 GTCAGCCTTCTACTTGAGCAAAATATTGATGTATCTTCTCAAGATCTATCTGGACAGACG 1020

Qy   1021 gccaaagagtatgctgtttctagtcatcatcatgtaatttgccagttacttttctgactac 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 GCCAGAGAGTATGCTGTTTCTAGTCATCATCATGTAATTTGCCAGTTACTTTTCTGACTAC 1080

Qy   1081 aaagaaaaacagatgctaaaaatctcttctgaaaacagcaatccagaa 1128
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 AAAGAAAAACAGATGCTAAAAATCTCTTCTGAAAACAGCAATCCAGAA 1128


RESULT   4
AX140884
LOCUS       AX140884     2000 bp    DNA             PAT       31-MAY-2001
DEFINITION  Sequence 374 from Patent WO0134802.
ACCESSION   AX140884
VERSION     AX140884.1  GI:14280987
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2000)
  AUTHORS   Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,
            Reed,S.G., Kalos,M.D., Retter,M.W., Stolk,J.A., Day,C.H.,
            Skeiky,Y.A. and Wang,A.
  TITLE     Compositions and methods for the therapy and diagnosis of prostate
            cancer
  JOURNAL   Patent: WO 0134802-A 374 17-MAY-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..2000
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      698 a    388 c    489 g    425 t
ORIGIN


  Query Match             97.7%;  Score 1128;  DB 6;  Length 2000;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1128;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 atggtggttgaggttgattccatgccggctgcctcttctgtgaagaagccatttggtctc 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ATGGTGGTTGAGGTTGATTCCATGCCGGCTGCCTCTTCTGTGAAGAAGCCATTTGGTCTC 60

Qy      61 aggaagcaagatgggcaagtggtgctgccgttgcttcccctgctgcagggaggagccggcaag 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 AGGAGCAAGATGGGCAAGTGGTGCTGCCGTTGCTTCCCCTGCTGCAGGGAGGAGCCGGCAAG 120

Qy     121 aggcaacgtgggcacttctggagaccacgacgactctgctatgaagacactcaggaagcaag 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 AGCAACGTGGGCACTTCTGGAGACCACGACGACTCTGCTATGAAGACACTCAGGAGCAAG 180

Qy     181 atgggcaagtggtgccgccgcactgcttccccctgctgcagggggagtgccaagagcaacgtg 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 ATGGGCAAGTGGTGCCGCCGCACTGCTTCCCCTGCTGCAGGGGGAGTGCCAAGAGCAACGTG 240

Qy     241 aggcgcttctggagaccacgacgactctgctatgaagacactcaggaacaagatgtggccaag 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 GGCGCTTCTGGAGACCACGACGACTCTGCTATGAAGACACTCAGGAACAACATGGGCAAG 300

Qy     301 tgatgctgccgctgcttccccctgctgcagggggagccggcaagagcaagggtgagccgcttgg 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 TGGTGCTGCCGCTGCTTCCCCTGCTGCAGGGGGAGCCGGCAAGAGCAAGGGTGAGCCGCTTGG 360

Qy     361 ggagactacgatgacagtgccttcatggagcccaggtaccacgtccgtggagaagatctg 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 GGAGACTACGATGACAGTGCCTTCATGGAGCCCAGGTACCACGTCCGTGGAGAAGATCTG 420

Qy     421 gacaagctccacagagctggcctggtgggggtaaagtcccccagaaggatctcatcgtcatg 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GACAAGCTCCACAGAGCTGGCCTGGTGGGGTAAAGTCCCCAGAAGGATCTCATCGTCATG 480

Qy     481 ctcaggcgacactgacgtgaacaagaagggacaagcaaaacaggactgctctacatctgccc 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 CTCAGGGACACTGACGTGAACAAGAAGGACAAGCAAAACAGGACTGCTCTACATCTGCCC 540

Qy     541 tctgccaatgggaattcagaagtagtaaaactcctgctggacagacgacgatgtcaacttaat 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 TCTGCCAATGGGAATTCAGAAGTAGTAAAACTCCTGCTGGACAGACGATGTCAACTTAAT 600

Qy     601 gtccttgacaacaaaaagaggacagctctgataaaggccgtacaatgccaggaagatgaa 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 GTCCTTGACAACAAAAAGAGGACAGCTCTGATAAAGGCCGTACAATGCCAGGAAGATGAA 660

Qy     661 tgtgcgttaatagttgctggaacatggcactgatccaaatattccagatgagtatggaaat 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 TGTGCGTTAATAGTTGCTGGAACATGGCACTGATCCAAATATTCCAGATGAGTATGGAAAT 720

Qy     721 accactctgccactacgctatctataatgaagataaattaatggccaaagcactgctctta 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 ACCACTCTGCCACTACGCTATCTATAATGAAGATAAATTAATGGCCAAAGCACTGCTCTTA 780

Qy     781 tatggtgctgatatcgaatcaaaaaacaagcatggcctcacacgcactgttacttggtgta 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 TATGGTGCTGATATCGAATCAAAAAACAAGCATGGCCTCACACGCACTGTTACTTGGTGTA 840

Qy     841 catgagcaaaaaacagcaagtcgtgaaattttaatcaagaaaaaagcgaatttaaatgca 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 CATGAGCAAAAACAGCAAGTCGTGAAATTTTAATCAAGAAAAAAGCGAATTTAAATGCA 900

Qy     901 ctggatagatatggaaggactgctctcatacttgctgtatgttgtggatcagcaagtata 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 CTGGATAGATATGGAAGGACTGCTCTCATACTTGCTGIATGTTGTGGATCAGCAAGTATA 96

Qy     961 gtcagccttctacttgagcaaaatattgatgtatcttctcaagatctatctggacagacg 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 GTCAGCCTTCTACTTGAGCAAAATATTGATGTATCTTCTCAAGATCTATCTGGACAGACG 1020

Qy    1021 gccaaagagtatgctgtttctagtcatcatcatgtaatttgccagttacttttctgactac 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 GCCAGAGAGTATGCTGTTTCTAGTCATCATCATGTAATTTGCCAGTTACTTTTCTGACTAC 1080

Qy    1081 aaagaaaaacagatgctaaaaatctcttctgaaaacagcaatccagaa 1128
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 AAAGAAAAACAGATGCTAAAAATCTCTTCTGAAAACAGCAATCCAGAA 1128


RESULT   5
AX106594
LOCUS       AX106594     2040 bp    DNA             PAT       30-APR-2001
DEFINITION  Sequence 375 from Patent WO0125272.
ACCESSION   AX106594
VERSION     AX106594.1  GI:13922265
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2040)
  AUTHORS   Xu,J., Skeiky,Y.A., Reed,S.G. and Cheever,M.A.
  TITLE     Compositions and methods for therapy and diagnosis of prostate
            cancer
  JOURNAL   Patent: WO 0125272-A 375 12-APR-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..2040
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      716 a    392 c    500 g    432 t
```

```
ORIGIN

  Query Match             97.7%;  Score 1128;  DB 6;  Length 2040;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1128;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 atggtggttgaggttgattccatgccggctgcctcttctgtgaagaagccatttggtctc 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGTGGTTGAGGTTGATTCCATGCCGGCTGCCTCTTCTGTGAAGAAGCCATTTGGTCTC 60

Qy         61 aggagcaagatgggcaagtggtgctgccgttgcttcccctgctgcagggagagcggcaag 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGGAGCAAGATGGGCAAGTGGTGCTGCCGTTGCTTCCCCTGCTGCAGGGAGAGCGGCAAG 120

Qy        121 agcaacgtgggcacttctggagaccacgacgactctgctatgaagacactcaggagcaag 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGCAACGTGGGCACTTCTGGAGACCACGACGACTCTGCTATGAAGACACTCAGGAGCAAG 180

Qy        181 atgggcaagtggtgccgccactgcttcccctgctgcagggggagtggcaagagcaacgtg 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATGGGCAAGTGGTGCCGCCACTGCTTCCCCTGCTGCAGGGGGAGTGGCAAGAGCAACGTG 240

Qy        241 ggcgcttctggagaccacgacgactctgctatgaagacactcaggaacaagatgggcaag 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GGCGCTTCTGGAGACCACGACGACTCTGCTATGAAGACACTCAGGAACAAGATGGGCAAG 300

Qy        301 tggtgctgccactgcttcccctgctgcagggggagcggcaagagcaaggtgggcgcttgg 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TGGTGCTGCCACTGCTTCCCCTGCTGCAGGGGGAGCGGCAAGAGCAAGGTGGGCGCTTGG 360

Qy        361 ggagactacgatgacagtgccttcatggagcccaggtaccacgtccgtggagaagatctg 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GGAGACTACGATGACAGTGCCTTCATGGAGCCCAGGTACCACGTCCGTGGAGAAGATCTG 420

Qy        421 gacaagctccacagagctgcctggtggggtaaagtccccagaaaggatctcatcgtcatg 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GACAAGCTCCACAGAGCTGCCTGGTGGGGTAAAGTCCCCAGAAAGGATCTCATCGTCATG 480

Qy        481 ctcagggacactgacgtgaacaagaaggacaagcaaaagaggactgctctacatctggcc 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTCAGGGACACTGACGTGAACAAGAAGGACAAGCAAAAGAGGACTGCTCTACATCTGGCC 540

Qy        541 tctgccaatgggaattcagaagtagtaaaactcctgctggacagacgatgtcaacttaat 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TCTGCCAATGGGAATTCAGAAGTAGTAAAACTCCTGCTGGACAGACGATGTCAACTTAAT 600

Qy        601 gtccttgacaacaaaaagaggacagctctgataaaggccgtacaatgccaggaagatgaa 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GTCCTTGACAACAAAAAGAGGACAGCTCTGATAAAGGCCGTACAATGCCAGGAAGATGAA 660

Qy        661 tgtgcgttaatgttgctggaacatggcactgatccaaatattccagatgagtatggaaat 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGTGCGTTAATGTTGCTGGAACATGGCACTGATCCAAATATTCCAGATGAGTATGGAAAT 720

Qy        721 accactctgcactacgctatctataatgaagataaattaatggccaaagcactgctctta 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ACCACTCTGCACTACGCTATCTATAATGAAGATAAATTAATGGCCAAAGCACTGCTCTTA 780

Qy        781 tatggtgctgatatcgaatcaaaaaacaagcatggcctcacaccactgttacttggtgta 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TATGGTGCTGATATCGAATCAAAAAACAAGCATGGCCTCACACCACTGTTACTTGGTGTA 840

Qy        841 catgagcaaaaacagcaagtcgtgaaatttttaatcaagaaaaaagcgaatttaaatgca 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CATGAGCAAAAACAGCAAGTCGTGAAATTTTTAATCAAGAAAAAAGCGAATTTAAATGCA 900

Qy        901 ctggatagatatggaaggactgctctcatacttgctgtatgttgtggatcagcaagtata 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CTGGATAGATATGGAAGGACTGCTCTCATACTTGCTGTATGTTGTGGATCAGCAAGTATA 960

Qy        961 gtcagccttctacttgagcaaaatattgatgtatcttctcaagatctatctggacagacg 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GTCAGCCTTCTACTTGAGCAAAATATTGATGTATCTTCTCAAGATCTATCTGGACAGACG 1020

Qy       1021 gccagagagtatgctgtttctagtcatcatcatgtaatttgccagttactttctgactac 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GCCAGAGAGTATGCTGTTTCTAGTCATCATCATGTAATTTGCCAGTTACTTTCTGACTAC 1080

Qy       1081 aaagaaaaacagatgctaaaaatctcttctgaaaacagcaatccagaa 1128
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AAAGAAAAACAGATGCTAAAAATCTCTTCTGAAAACAGCAATCCAGAA 1128


RESULT 6
AX140885
LOCUS       AX140885     2040 bp    DNA             PAT       31-MAY-2001
DEFINITION  Sequence 475 from Patent WO0134802.
ACCESSION   AX140885
VERSION     AX140885.1  GI:14280988
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2040)
  AUTHORS   Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,
            Reed,S.G., Kalos,M.D., Retter,M.W., Stolk,J.A., Day,C.H.,
            Skeiky,Y.A. and Wang,A.
  TITLE     Compositions and methods for the therapy and diagnosis of prostate
            cancer
  JOURNAL   Patent: WO 0134802-A 475 17-MAY-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..2040
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      716 a    492 c    500 g    432 t
ORIGIN

  Query Match             97.7%;  Score 1128;  DB 6;  Length 2040;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1128;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 atggtggttgaggttgattccatgccggctgcctcttctgtgaagaagccatttggtctc 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGTGGTTGAGGTTGATTCCATGCCGGCTGCCTCTTCTGTGAAGAAGCCATTTGGTCTC 60

Qy         61 aggagcaagatgggcaagtggtgctgccgttgcttcccctgctgcagggagagcggcaag 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGGAGCAAGATGGGCAAGTGGTGCTGCCGTTGCTTCCCCTGCTGCAGGGAGAGCGGCAAG 120

Qy        121 agcaacgtgggcacttctggagaccacgacgactctgctatgaagacactcaggagcaag 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGCAACGTGGGCACTTCTGGAGACCACGACGACTCTGCTATGAAGACACTCAGGAGCAAG 180

Qy        181 atgggcaagtggtgccgccactgcttcccctgctgcagggggagtggcaagagcaacgtg 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATGGGCAAGTGGTGCCGCCACTGCTTCCCCTGCTGCAGGGGGAGTGGCAAGAGCAACGTG 240

Qy        241 ggcgcttctggagaccacgacgactctgctatgaagacactcaggaacaagatgggcaag 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GGCGCTTCTGGAGACCACGACGACTCTGCTATGAAGACACTCAGGAACAAGATGGGCAAG 300

Qy        301 tggtgctgccactgcttcccctgctgcagggggagcggcaagagcaaggtgggcgcttgg 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TGGTGCTGCCACTGCTTCCCCTGCTGCAGGGGGAGCGGCAAGAGCAAGGTGGGCGCTTGG 360

Qy        361 ggagactacgatgacagtgccttcatggagcccaggtaccacgtccgtggagaagatctg 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GGAGACTACGATGACAGTGCCTTCATGGAGCCCAGGTACCACGTCCGTGGAGAAGATCTG 420

Qy        421 gacaagctccacagagctgcctggtggggtaaagtccccagaaaggatctcatcgtcatg 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GACAAGCTCCACAGAGCTGCCTGGTGGGGTAAAGTCCCCAGAAAGGATCTCATCGTCATG 480
```

```
Qy    481 ctcagcgacactgacgtgaacaacaagcacaagcaaaacaggactgctctacatctgccc 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CTCAGGGACACTGACGTGAACAACAAGCACAAGCAAAACAGGACTGCTCTACATCTGGCC 540

Qy    541 tctgccaatgggaattcagaagtagtaaaactcctgctggacagacgatgtcaacttaat 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 TCTGCCAATGGGAATTCAGAAGTAGTAAAACTCCTGCTGGACAGACGATGTCAACTTAAT 600

Qy    601 gtccttgacaacaaaaagagggacagctctgataaaggccgtacaatgccaggaagatgaa 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 GTCCTTGACAACAAAAAGAGGACAGCTCTGATAAAGGCCGTACAATGCCAGGAAGATGAA 660

Qy    661 tgtgcgttaatgttgctggaacatggcactgatccaaatattccagatgagtatggaaat 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 TGTGCGTTAATGTTGCTGGAACATGGCACTGATCCAAATATTCCAGATGAGTATGGAAAT 720

Qy    721 accactctgcactacgctatctataatgaagataaattaatggccaaagcactgctctta 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 ACCACTCTGCACTACGCTATCTATAATGAAGATAAATTAATGGCCAAAGCACTGCTCTTA 780

Qy    781 tatggtgctgatatcgaatcaaaaaacaagcatggcctcacaccactgttacttggtgta 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 TATGGTGCTGATATCGAATCAAAAAACAAGCATGGCCTCACACCACTGTTACTTGGTGTA 840

Qy    841 catgagcaaaaacagcaagtcgtgaaatttttaatcaagaaaaaagcgaatttaaatgca 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CATGAGCAAAAACAGCAAGTCGTGAAATTTTTAATCAAGAAAAAAGCGAATTTAAATGCA 900

Qy    901 ctggatagatatgaaagactgctctcatacttgctgtatgttgtggatcagcaagtata 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 CTGGATAGATATGGAAGGACTGCTCTCATACTTGCTGTATGTTGTGGATCAGCAAGTATA 960

Qy    961 gtcagccttctacttgagcaaaatattgatgtatcttctcaagatctatctggacagacg 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 GTCAGCCTTCTACTTGAGCAAAATATTGATGTATCTTCTCAAGATCTATCTGGACAGACG 1020

Qy   1021 gccagagagtatgctgttttctagtcatcatcatgtaatttgccagttacttttctgactac 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 GCCAGAGAGTATGCTGTTTTCTAGTCATCATCATGTAATTTGCCAGTTACTTTTCTGACTAC 1080

Qy   1081 aaagaaaaacagatgctaaaaatctcttctgaaaacagcaatccagaa 1128
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 AAAGAAAAACAGATGCTAAAAATCTCTTCTGAAAACAGCAATCCAGAA 1128


RESULT   7
AX106587
LOCUS       AX106587     1512 bp    DNA             PAT       30-APR-2001
DEFINITION  Sequence 368 from Patent WO0125272.
ACCESSION   AX106587
VERSION     AX106587.1  GI:13922258
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1512)
  AUTHORS   Xu,J., Skeiky,Y.A., Reed,S.G. and Cheever,M.A.
  TITLE     Compositions and methods for therapy and diagnosis of prostate
            cancer
  JOURNAL   Patent: WO 0125272-A 368 12 APR-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .1512
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      406 a    301 c    393 g    399 t     13 others
ORIGIN


  Query Match             42.5%;  Score 491;  DB 6;  Length 1512;
  Best Local Similarity   99.4%;  Pred. No. 2.8e-248;
  Matches  691;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy     34 tcttctgtgaagaagccattggctctcaggagcaagatgggcaagtggtgctgccgttgc 93
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    256 TCTTCTGTGAAGAAGCCATTGGTCTCAGGAGCAAGATGGGCAAGTGGTGCTGCCGTTGC 315

Qy     94 ttccctgctgcaggagagcggcaagagcaacgtgggcacttctggagaccacgacgac 153
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    316 TTCCCCTGCTGCAGGGAGAGCGGCAAGAGCAACGTGGGCACTTCTGGAGACCACGACGAC 375

Qy    154 tctgctatgaagacactcaggagcaagatgggcaagtggtgccgccactgcttccctgc 213
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    376 TCTGCTATGAAGACACTCAGGAGCAAGATGGGCAAGTGGTGCCGCCACTGCTTCCCCTGC 435

Qy    214 tgcagggagagtggcaagagcaacgtgggcgcttctggagaccacgacgactctgctatg 273
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    436 TGCAGGGGAGTGGCAAGAGCAACGTGGGCGCTTCTGGAGACCACGACGAYTCTGCTATG 495

Qy    274 aagacactcaggaacaagatgggcaagtggtgctgccactgcttccctgctgcagggggg 333
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    496 AAGACACTCAGGAACAAGATGGGCAAGTGGTGCTGCCACTGCTTCCCCTGCTGCAGGGGG 555

Qy    334 agcggcaagagcaagggtgggcgcttgggagactacgatgacaagtgccttcatggagccc 393
          ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    556 AGCRGCAAGAGCAAGGTGGGCGCTTGGGAGACTACGATGACAGTGCCTTCATGGAGCCC 615

Qy    394 aggtaccacgtccgtgagaagatctggacaagctccacacagctgccgtggtggtaaa 453
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    616 AGGTACCACGTCCGTGGAGAAGATCTGGACAAGCTCCACACAGCTGCCGTGGTGGTAAA 675

Qy    454 gtcccagaaagaatctcatcgtcatgctcaggacactgacgtggaacaagaagaggacaag 513
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    676 GTCCCAGAAAGGATCTCATCGTCATGCTCAGGACACTGACGTGGAACAAGAAGGACAAG 735

Qy    514 caaaagaggactgctctacatctggcctctgccaatgggaattcagaagtagtaaaactc 573
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    736 CAAAAGAGGACTGCTCTACATCTGGCCTCTGCCAATGGGAATTCAGAAGTAGTAAAACTC 795

Qy    574 ctgctggacagacgatgtcaacttaatgtccttgacaacaaaaagaggacagctctgata 633
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    796 STGCTGGACAGACGATGTCAACTTAATGTCCTTGACAACAAAAAGAGGACAGCTCTGAYA 855

Qy    634 aaggccgtacaatgccaggaagatgaatgtgcgttaatgttgctggaacatggcactgat 693
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    856 AAGGCCGTACAATGCCAGGAAGATGAATGTGCGTTAATGTTGCTGGAACATGGCACTGAT 915

Qy    694 ccaaatattccagatgagtatggaaataccactct 728
          |||||||||||||||||||||||||||||||||||
Db    916 CCAAATATTCCAGATGAGTATGGAAATACCACTCT 950


RESULT   8
AX140878
LOCUS       AX140878     1512 bp    DNA             PAT       31-MAY-2001
DEFINITION  Sequence 368 from Patent WO0134802.
ACCESSION   AX140878
VERSION     AX140878.1  GI:14280981
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1512)
  AUTHORS   Xu,J., Dillon,P.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,
            Reed,S.G., Kalos,M.D., Retter,M.W., Stolk,J.A., Day,C.H.,
            Skeiky,Y.A. and Wang,A.
  TITLE     Compositions and methods for the therapy and diagnosis of prostate
            cancer
  JOURNAL   Patent: WO 0134802-A 368 17-MAY-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .1512
                     /organism="Homo sapiens"
```

```
                    /db_xref="taxon:9606"
BASE COUNT      40?? a    301 c    393 g    399 t     13 others
ORIGIN

  Query Match             42.5%;  Score 491;  DB 6;  Length 1512;
  Best Local Similarity   99.4%;  Pred. No. 2.8e-248;
  Matches  ???;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy         34 tctttctgtgaagaagccatttggtctcaggagcaagatgggcaagtggtgctgccgttgc 93
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        256 TCTTCTGTGAAGAAGCCATTTGGTCTCAGGAGCAAGATGGGCAAGTGGTGCTGCCGTTGC 315

Qy         94 ttcccctgctgcaggaagacggcaagagcaacgtgggcacttctggagaccacgacgac 153
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        316 TTCCCCTGCTGCAGGGAGACGGCAAGAGCAACGTGGGCACTTCTGGAGACCACGACGAC 375

Qy        154 tctgctatgaagactctcaggagcaagatgggcaagtggtgccgccactgcttcccctgc 213
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        376 TCTGCTATGAAGACACTCAGGAGCAAGATGGGCAAGTGGTGCCGCCACTGCTTCCCCTGC 435

Qy        214 tgcagggggagtggcaagagcaacgtgggcgcttctggagaccacgacgactctgctatg 273
              ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db        436 TGCAGGGGGAGTGGCAAGAGCAACGTGGGCGCTTCTGGAGACCACGACGAYTCTGCTATG 495

Qy        274 aagacactcaggaacaagatgggcaagtggtgctgccactgcttcccctgctgcaggggg 333
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        496 AAGACACTCAGGAACAAGATGGGCAAGTGGTGCTGCCACTGCTTCCCCTGCTGCAGGGGG 555

Qy        334 agcggcaagagcaaggtgggcgcttggggagactacgatgacagtgccttcatggagccc 393
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        556 AGCGGTAAGAGCAAGGTGGGCGCTTGGGGAGACTACGATGACAGTGCCTTCATGGAGCCC 615

Qy        394 aggtaccacgtccgtggagaagatctggacaagctccacagagctgcctggtggggtaaa 453
              ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        616 AGGTACCACGTCCGTGGAGAAGATCTGGACAAGCTCCACAGAGCTGCCTGGTGGGGTAAA 675

Qy        454 gtccccagaaaggatctcatcgtcatgctcagggacactgacgtgaacaagaaggacaag 513
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        676 GTCCCCAGAAAGGATCTCATCGTCATGCTCAGGGACACTGACGTGAACAAGAAGGACAAG 735

Qy        514 caaaagaggactgctctacatctggcctctgccaatgggaattcagaagtagtaaaactc 573
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        736 CAAAAGAGGACTGCTCTACATCTGGCCTCTGCCAATGGGAATTCAGAAGTAGTAAAACTC 795

Qy        574 ctgctggacagacgatgtcaacttaatgtccttgacaacaaaaagaggacagctctgata 633
              |  ||||||||||||||||||||||||||| |||||||||||||||||||||||||||  |
Db        796 STGCTGGACAGACGATGTCAACTTAATGTCCTTGACAACAAAAAGAGGACAGCTCTGAYA 855

Qy        634 aaggccgtacaatgccaggaagatgaatgtgcgttaatgttgctggaacatggcactgat 693
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        856 AAGGCCGTACAATGCCAGGAAGATGAATGTGCGTTAATGTTGCTGGAACATGGCACTGAT 915

Qy        694 ccaaatattccagatgagtatggaaatacccactct 728
              ||||||||||||||||||||||||||||||||||||
Db        916 CCAAATATTCCAGATGAGTATGGAAATACCCACTCT 950


RESULT 5
AC018804/c
LOCUS       AC018804    199668 bp    DNA             HTG       19-APR-2001
DEFINITION  Homo sapiens chromosome 2 clone RP11-397H17, WORKING DRAFT
            SEQUENCE, 7 unordered pieces.
ACCESSION   AC018804
VERSION     AC018804.8  GI:13677145
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 199668)
  AUTHORS   Waterston,R.H.
  TITLE     The sequence of Homo sapiens clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 199668)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-JAN-2000) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     On Apr 19, 2001 this sequence version replaced gi:13270700.
            ----------- Genome Center
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site: http://genome.wustl.edu/gsc/index.shtml
            ----------- Project Information
            Center project name: H_NH0397H17
            ----------- Summary Statistics
            Sequencing vector: M13; 43%
            Sequencing vector: plasmid; 44%
            Chemistry: Dye primer ET; 43% of reads
            Chemistry: Dye terminator Big Dye; 44% of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 199883 bases at least Q40
            Consensus quality: 196289 bases at least Q30
            Consensus quality: 196285 bases at least Q20
            Insert size: 192000; agarose-fp
            Insert size: 200232; sum-of-contigs
            Quality coverage: 6.33 in Q20 bases; agarose-fp
            Quality coverage: 6.46 in Q20 bases; sum-of-contigs
            ---------------
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 7 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1    2205: contig of 2205 bp in length
            *     2206    2305: gap of unknown length
            *     2306    6011: contig of 3706 bp in length
            *     6012    6111: gap of unknown length
            *     6112   11626: contig of 5515 bp in length
            *    11627   11726: gap of unknown length
            *    11727   36054: contig of 24328 bp in length
            *    36055   36154: gap of unknown length
            *    36155   58300: contig of 22146 bp in length
            *    58301   58400: gap of unknown length
            *    58401  121289: contig of 62889 bp in length
            *   121290  121389: gap of unknown length
            *   121390  199668: contig of 78279 bp in length.
FEATURES             Location/Qualifiers
     source          1..199668
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="2"
                     /clone="RP11-397H17"
     misc_feature    1..2205
                     /note="assembly_name:Contig18"
     misc_feature    2306..6011
                     /note="assembly_name:Contig16"
     misc_feature    6112..11626
                     /note="assembly_name:Contig20"
     misc_feature    11727..36054
                     /note="assembly_name:Contig21
                     clone_end:T7
                     vector_side:left"
     misc_feature    36155..58300
                     /note="assembly_name:Contig22
                     clone_end:SP6
                     vector_side:right"
     misc_feature    58401..121289
                     /note="assembly_name:Contig23"
     misc_feature    121390..199668
```

```
                     /db_xref "assembly_name:Contig24"
BASE COUNT    53956 a   46083 c   46485 g   52543 t     601 others
ORIGIN


  Query Match             39.2%;  Score 453;  DB 2;  Length 199668;
  Best Local Similarity   99.8%;  Pred. No. 2.8e-228;
  Matches  503;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 atggtggttgaggttgattccatgccggctgccttcttctgtgaagaagccatttggtctc 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  110826 ATGGTGGTTGAGGTTGATTCCATGCCGGCTGCCTTCTTCTGTGAAGAAGCCATTTGGTCTC 110767

Qy      61 aggaagcaagatgtggcaagtggtgctgccgttgcttccccctgctgcaggggagagccggcaaag 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  110766 AGGAAGCAAGATGGGCAAGTGGTGCTGCCGTTGCTTCCCCCTGCTGCAGGGAGAGCCGGCAAG 110707

Qy     121 agcaacgtgaggcacttctgggagaccacgaacgactctgctatgaagacactcaggagcaag 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  110706 AGCAACGTGGGCACTTCTGGAGACCACGACGACTCTGCTATGAAGACACTCAGGAGCAAG 110647

Qy     181 atgggcaagtggtgccgccactgcttccccctgctgcaggggagtggcaagagcaacgtg 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  110646 ATGGGCAAGTGGTGCCGCCACTGCTTCCCCCTGCTGCAGGGGAGTGGCAAGAGCAACGTG 110587

Qy     241 ggcgcttctgaagaccacgacgactctgctatgaagacactcaggaacaagatgggcaag 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  110586 GGCGCTTCTGGAGACCACGACGACTCTGCTATGAAGACACTCAGGAACAAGATGGGCAAG 110527

Qy     301 tggtgctgccactgcttccccctgctgcaggggagcaggcaagagcaagtgggcgcttgg 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  110526 TGGTGCTGCCACTGCTTCCCCCTGCTGCAGGGGAGCAGCAAGAGCAAGGTGGGCGCTTGG 110467

Qy     361 ggagactacgatgacagtgccttcatggagcccaggtaccacgtccgtgaagaagatctg 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  110466 GGAGACTACGATGACAGTGCCTTCATGGAGCCCAGGTACCACGTCCGTGAGAAGATCTG 110407

Qy     421 [illegible] 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  110406 GACAAGCTCCACAGAGCTGCCTGGTGGGGTAAAGTCCCCAGAAAGGATCTCATCGTCATG 110347

Qy     481 ctcaggacactgacgtgaacaag 504
           |||||||||||||||||||||||
Db  110346 CTCAGGGACACTGACGTGAACAAG 110323


RESULT  10
AX106588
LOCUS       AX106588     1853 bp    DNA             PAT       30-APR-2001
DEFINITION  Sequence 369 from Patent WO0125272.
ACCESSION   AX106588
VERSION     AX106588.1  GI:14922259
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1853)
  AUTHORS   Xu,J., Skeiky,Y.A., Reed,S.G. and Cheever,M.A.
  TITLE     Compositions and methods for therapy and diagnosis of prostate
            cancer
  JOURNAL   Patent: WO 0125272-A 369 12-APR-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .1853
                     /organism="Homo sapiens"
                     /db_xref "taxon:9606"
BASE COUNT      521 a    381 c    492 g    431 t     28 others
ORIGIN


  Query Match             21.8%;  Score 252;  DB 6;  Length 1853;
  Best Local Similarity   99.7%;  Pred. No. 7e-122;
  Matches  302;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      34 tcttctgtgaagaagccatttggtctcaggagcaagatgggcaagtggtgctgccgttgc 93
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     256 TCTTCTGTGAAGAAGCCATTTGGTCTCAGGAGCAAGATGGGCAAGTGGTGCTGCCGTTGC 315

Qy      94 ttccccctgctgcagggagagccggcaagagcaacgtgggcacttctggagaccacgacgac 153
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     316 TTCCCCTGCTGCAGGGAGAGCCGGCAAGAGCAACGTGGGCACTTCTGGAGACCACGACGAC 375

Qy     154 tctgctatgaagacactcaggagcaagatgggcaagtggtgccgccactgcttccccctgc 213
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     376 TCTGCTATGAAGACACTCAGGAGCAAGATGGGCAAGTGGTGCCGCCACTGCTTCCCCTGC 435

Qy     214 tgcaggggagtggcaagagcaacgtgggcgcttctggaagaccacgacgactctgctatg 273
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     436 TGCAGGGGAGTGGCAAGAGCAACGTGGGCGCTTCTGGAGACCACGACGAYTCTGCTATG 495

Qy     274 aagacactcaggaacaagatgggcaagtggtgctgccactgcttccccctgctgcagggggg 333
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     496 AAGACACTCAGGAACAAGATGGGCAAGTGGTGCTGCCACTGCTTCCCCTGCTGCAGGGGG 555

Qy     334 agc 336
           |||
Db     556 AGC 558


RESULT  11
AX140879
LOCUS       AX140879     1853 bp    DNA             PAT       31-MAY-2001
DEFINITION  Sequence 369 from Patent WO0134802.
ACCESSION   AX140879
VERSION     AX140879.1  GI:14280982
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1853)
  AUTHORS   Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,
            Reed,S.G., Kalos,M.D., Retter,M.W., Stolk,J.A., Day,C.H.,
            Skeiky,Y.A. and Wang,A.
  TITLE     Compositions and methods for the therapy and diagnosis of prostate
            cancer
  JOURNAL   Patent: WO 0134802-A 369 17-MAY-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .1853
                     /organism "Homo sapiens"
                     /db_xref "taxon:9606"
BASE COUNT      521 a    381 c    492 g    431 t     28 others
ORIGIN


  Query Match             21.8%;  Score 252;  DB 6;  Length 1853;
  Best Local Similarity   99.7%;  Pred. No. 7e-122;
  Matches  302;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      34 tcttctgtgaagaagccatttggtctcaggagcaagatgggcaagtggtgctgccgttgc 93
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     256 TCTTCTGTGAAGAAGCCATTTGGTCTCAGGAGCAAGATGGGCAAGTGGTGCTGCCGTTGC 315

Qy      94 ttccccctgctgcagggagagccggcaagagcaacgtgggcacttctggagaccacgacgac 153
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     316 TTCCCCTGCTGCAGGGAGAGCCGGCAAGAGCAACGTGGGCACTTCTGGAGACCACGACGAC 375

Qy     154 tctgctatgaagacactcaggagcaagatgggcaagtggtgccgccactgcttccccctgc 213
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     376 TCTGCTATGAAGACACTCAGGAGCAAGATGGGCAAGTGGTGCCGCCACTGCTTCCCCTGC 435

Qy     214 tgcagggagtggcaagagcaacgtgggcgcttctggaagaccacgacgactctgctatg 273
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     436 TGCAGGGGAGTGGCAAGAGCAACGTGGGCGCTTCTGGAGACCACGACGAYTCTGCTATG 495
```

```
Qy      274 aagacactcaggaacaagatgggcaagtggtgctgccactgcttcccctgctgcagggg 333
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      496 AAGACACTCAGGAACAAGATGGGCAAGTGGTGCTGCCACTGCTTCCCCTGCTGCAGGGG 555

Qy      334 agc 336
            |||
Db      556 AGC 558


RESULT 12
AC013269/c
LOCUS       AC013269    115727 bp    DNA             HTG       22 MAR 2001
DEFINITION  Homo sapiens chromosome UNK clone RP11 356I10, WORKING DRAFT
            SEQUENCE, 9 unordered pieces.
ACCESSION   AC013269
VERSION     AC013269.9  GI:13431162
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 115727)
  AUTHORS   Waterston,R.H.
  TITLE     The sequence of Homo sapiens clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 115727)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-NOV-1999) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     On Mar 22, 2001 this sequence version replaced gi:11991325.

                         Genome Center
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site: http://genome.wustl.edu/gsc/index.shtml
                         Project Information
            Center project name: H_NH0356I10
                         Summary Statistics
            Sequencing vector: M13; 45%
            Sequencing vector: plasmid; 42%
            Chemistry: Dye-primer ET; 45% of reads
            Chemistry: Dye-terminator Big Dye; 42% of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 112722 bases at least Q40
            Consensus quality: 113531 bases at least Q30
            Consensus quality: 113947 bases at least Q20
            Insert size: 122000; agarose-fp
            Insert size: 114927; sum of contigs
            Quality coverage: 6.17 in Q20 bases; agarose-fp
            Quality coverage: 8.31 in Q20 bases; sum of contigs

            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 9 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *          1      1137: contig of 1137 bp in length
            *       1138      1237: gap of unknown length
            *       1238      2559: contig of 1322 bp in length
            *       2560      2659: gap of unknown length
            *       2660      4795: contig of 2136 bp in length
            *       4796      4895: gap of unknown length
            *       4896      8967: contig of 4072 bp in length
            *       8968      9067: gap of unknown length
            *       9068     14353: contig of 5286 bp in length
            *      14354     14453: gap of unknown length
            *      14454     20926: contig of 6473 bp in length
            *      20927     21026: gap of unknown length
            *      21027     30828: contig of 9802 bp in length
            *      30829     30928: gap of unknown length
            *      30929     66784: contig of 35856 bp in length
            *      66785     66884: gap of unknown length
            *      66885    115727: contig of 48843 bp in length
FEATURES             Location/Qualifiers
     source          1..115727
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="UNK"
                     /clone="RP11-356I10"
     misc_feature    1..1137
                     /note="assembly_name:Contig12"
     misc_feature    1238..2559
                     /note="assembly_name:Contig13"
     misc_feature    2660..4795
                     /note="assembly_name:Contig14"
     misc_feature    4896..8967
                     /note="assembly_name:Contig15"
     misc_feature    9068..14353
                     /note="assembly_name:Contig16"
     misc_feature    14454..20926
                     /note="assembly_name:Contig17"
     misc_feature    21027..30828
                     /note="assembly_name:Contig18"
     misc_feature    30929..66784
                     /note="assembly_name:Contig19"
     misc_feature    66885..115727
                     /note="assembly_name:Contig20"
BASE COUNT    31767 a  24959 c  25770 g  32438 t    805 others
ORIGIN

  Query Match             11.9%;  Score 137;  DB 2;  Length 115727;
  Best Local Similarity   100.0%;  Pred. No. 3.9e-61;
  Matches  137;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      919 actgctctcatacttgctgtatgttgtggatcagcaagtatagtcagccttctacttgag 978
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    23100 ACTGCTCTCATACTTGCTGTATGTTGTGGATCAGCAAGTATAGTCAGCCTTCTACTTGAG 23041

Qy      979 caaaatattgatgtatcttctcaagatctatctggacagacggccagagagtatgctgtt 1038
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    23040 CAAAATATTGATGTATCTTCTCAAGATCTATCTGGACAGACGGCCAGAGAGTATGCTGTT 22981

Qy     1039 tctagtcatcatcatgt 1055
            |||||||||||||||||
Db    22980 TCTAGTCATCATCATGT 22964


RESULT 13
AX141041
LOCUS       AX141041      879 bp    DNA             PAT       31-MAY-2001
DEFINITION  Sequence 531 from Patent WO0134802.
ACCESSION   AX141041
VERSION     AX141041.1  GI:14281098
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 879)
  AUTHORS   Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,
            Reed,S.G., Kalos,M.D., Retter,M.W., Stolk,J.A., Day,C.H.,
            Skeiky,Y.A. and Wang,A.
  TITLE     Compositions and methods for the therapy and diagnosis of prostate
            cancer
  JOURNAL   Patent: WO 0134802-A 531 17-MAY-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..879
                     /organism="Homo sapiens"
```

```
                     /db_xref="taxon:9606"
BASE COUNT      263 a    193 c    222 g    201 t
ORIGIN


  Query Match             10.4%;  Score 120;  DB 6;  Length 879;
  Best Local Similarity   99.1%;  Pred. No. 4.7e-52;
  Matches  220;  Conservative    0;  Mismatches    2;  Indels    0,  Gaps    0,

Qy     507 qqacaaqcaaaaaqaqqactqctctacatctggcctctgccaatgqqaattcaqaaqtagt 566
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     333 GGACAAGCAAAAGAGGACTGCTCTACATCTGGCCTCTGCCAATGGCAATTCAGAAGTAGT 392

Qy     567 aaaactcctgctqqacagacgatgtcaacttaatgtccttgacaacaaaaaqaqqacaqc 626
           ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     393 AAAACTCGTGCTGGACAGACGATGTCAACTTAATGTCCTTGACAACAAAAAGAGGACAGC 452

Qy     627 tctqataaaggccqtacaatqccaggaagatgaatgtgcgttaatgttgctggaacatgg 686
           ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     453 TCTGACAAAGCCCGTACAATGCCAGGAAGATGAATGTGCGTTAATGTTGCTGGAACATGG 512

Qy     687 cactgatccaaatattccagatgagtatggaaataccactct 728
           |||||||||||||||||||||||||||||||||||||||||||
Db     513 CACTGATCCAAATATTCCAGATGAGTATGGAAATACCACTCT 554


RESULT  14
AX106591
LOCUS       AX106591     1059 bp    DNA              PAT       30-APR-2001
DEFINITION  Sequence 372 from Patent WO0125272.
ACCESSION   AX106591
VERSION     AX106591.1  GI:14922262
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1059)
  AUTHORS   Xu,J., Skolky,Y.A., Reed,S.G. and Cheever,M.A.
  TITLE     Compositions and methods for therapy and diagnosis of prostate
            cancer
  JOURNAL   Patent: WO 0125272 A 372 12-APR-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .1059
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      325 a    204 c    232 g    286 t     12 others
ORIGIN


  Query Match             10.4%;  Score 120;  DB 6;  Length 1059;
  Best Local Similarity   99.1%;  Pred. No. 4.7e-52;
  Matches  220;  Conservative    0;  Mismatches    2;  Indels    0,  Gaps    0;

Qy     507 qqacaaqcaaaaaqaqqactqctctacatctggcctctgccaatgqqaattcaqaaqtagt 566
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     276 GGACAAGCAAAAGAGGACTGCTCTACATCTGGCCTCTGCCAATGGCAATTCAGAAGTAGT 335

Qy     567 aaaactcctgctqqacagacgatgtcaacttaatgtccttgacaacaaaaaqaqqacaqc 626
           ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     336 AAAACTCGTGCTGGACAGACGATGTCAACTTAATGTCCTTGACAACAAAAAGAGGACAGC 395

Qy     627 tctqataaaggccqtacaatqccaggaagatgaatgtgcgttaatgttgctggaacatgg 686
           ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     396 TCTGAYAAAGCCCGTACAATGCCAGGAAGATGAATGTGCGTTAATGTTGCTGGAACATGG 455

Qy     687 cactgatccaaatattccagatgagtatggaaataccactct 728
           |||||||||||||||||||||||||||||||||||||||||||
Db     456 CACTGATCCAAATATTCCAGATGAGTATGGAAATACCACTCT 497


RESULT  15
```

```
AX140882
LOCUS       AX140882     1059 bp    DNA              PAT       31-MAY-2001
DEFINITION  Sequence 372 from Patent WO0134802.
ACCESSION   AX140882
VERSION     AX140882.1  GI:14280985
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1059)
  AUTHORS   Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,
            Reed,S.G., Kalos,M.D., Retter,M.W., Stolk,J.A., Day,C.H.,
            Skolky,Y.A. and Wang,A.
  TITLE     Compositions and methods for the therapy and diagnosis of prostate
            cancer
  JOURNAL   Patent: WO 0134802-A 372 17-MAY-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .1059
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      325 a    204 c    232 g    286 t     12 others
ORIGIN


  Query Match             10.4%;  Score 120;  DB 6;  Length 1059;
  Best Local Similarity   99.1%;  Pred. No. 4.7e-52;
  Matches  220;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy     507 qqacaaqcaaaaaqaqqactqctctacatctggcctctgccaatgqqaattcaqaaqtagt 566
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     276 GGACAAGCAAAAGAGGACTGCTCTACATCTGGCCTCTGCCAATGGCAATTCAGAAGTAGT 335

Qy     567 aaaactcctgctqqacagacgatgtcaacttaatgtccttgacaacaaaaaqaqqacaqc 626
           ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     336 AAAACTCGTGCTGGACAGACGATGTCAACTTAATGTCCTTGACAACAAAAAGAGGACAGC 395

Qy     627 tctqataaaggccqtacaatqccaggaagatgaatgtgcgttaatgttgctggaacatgg 686
           ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     396 TCTGAYAAAGCCCGTACAATGCCAGGAAGATGAATGTGCGTTAATGTTGCTGGAACATGG 455

Qy     687 cactgatccaaatattccagatgagtatggaaataccactct 728
           |||||||||||||||||||||||||||||||||||||||||||
Db     456 CACTGATCCAAATATTCCAGATGAGTATGGAAATACCACTCT 497


Search completed: March  9, 2002, 04:14:51
Job time: 6799 sec
```

OM nucleic - nucleic search, using sw model

Run on: March 9, 2002, 00:37:16 ; Search time 77.78 Seconds
(without alignments)
[illegible] Million cell updates/sec

Title: US-09-699-295-301
Perfect score: 1155
Sequence: 1 atggtggttgaggttgattc.........caagaaccagaaataaataa 1155

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 951203 seqs, 1132389[illegible] residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Issued Patents NA:*
1: [illegible]_COMB.seq:*
2: [illegible]_COMB.seq:*
3: [illegible]_COMB.seq:*
4: [illegible]_COMB.seq:*
5: [illegible]_COMB.seq:*
6: [illegible].seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	829	71.8	1851	4	US-08-991-789A-291	Sequence 291, App
	2	[illegible]	[illegible]	451	4	US-08-991-789A-211	Sequence 211, App
	3	84.0	7.3	3454	4	US-09-082-059-1	Sequence 1, Appli
	4	82.6	7.2	1056	3	US-09-065-474-138	Sequence 138, App
c	5	82.6	7.2	1056	3	US-09-065-474-140	Sequence 140, App
	6	82.6	7.2	5235	1	US-09-031-485-35	Sequence 35, Appl
	7	82.6	7.2	5235	1	US-09-031-485-36	Sequence 36, Appl
	8	82.6	7.2	5235	1	US-08-847-429A-35	Sequence 35, Appl
c	9	82.6	7.2	5235	1	US-08-847-429A-36	Sequence 36, Appl
	10	82.6	7.2	5235	3	US-09-065-474-35	Sequence 35, Appl
	11	82.6	7.2	5235	3	US-09-065-474-36	Sequence 36, Appl
	12	82.6	7.2	5503	1	US-09-031-485-32	Sequence 32, Appl
	13	82.6	7.2	5503	1	US-09-031-485-34	Sequence 34, Appl
	14	82.6	7.2	5503	1	US-08-847-429A-32	Sequence 32, Appl
c	15	82.6	7.2	5503	1	US-08-847-429A-34	Sequence 34, Appl
	16	82.6	7.2	5503	3	US-09-065-474-32	Sequence 32, Appl
c	17	82.6	7.2	5503	3	US-09-065-474-34	Sequence 34, Appl
	18	63.2	5.5	906	1	US-09-031-485-40	Sequence 40, Appl
	19	63.2	5.5	906	1	US-09-031-485-41	Sequence 41, Appl
	20	63.2	5.5	906	1	US-08-847-429A-40	Sequence 40, Appl
c	21	63.2	5.5	906	1	US-08-847-429A-41	Sequence 41, Appl
	22	63.2	5.5	906	3	US-09-065-474-40	Sequence 40, Appl
c	23	63.2	5.5	906	3	US-09-065-474-41	Sequence 41, Appl
	24	63.2	5.5	908	1	US-09-031-485-37	Sequence 37, Appl
	25	63.2	5.5	908	1	US-09-031-485-39	Sequence 39, Appl
	26	63.2	5.5	908	1	US-08-847-429A-37	Sequence 37, Appl
c	27	63.2	5.5	908	1	US-08-847-429A-39	Sequence 39, Appl
	28	63.2	5.5	908	3	US-09-065-474-37	Sequence 37, Appl
c	29	63.2	5.5	908	3	US-09-065-474-39	Sequence 39, Appl
	30	62.2	5.4	4291	2	US-08-417-210A-80	Sequence 80, Appl
	31	62.2	5.4	4857	2	US-08-566-398-16	Sequence 16, Appl
	32	62.2	5.4	6628	3	US-08-815-809-3	Sequence 3, Appl
	33	62.2	5.4	6649	2	US-08-816-155B-5	Sequence 5, Appl
	34	62.2	5.4	6649	3	US-09-079-587-5	Sequence 5, Appl
	35	62.2	5.4	7091	2	US-08-658-665-40	Sequence 40, Appl
	36	62.2	5.4	7091	4	US-08-796-101-4	Sequence 4, Appl
	37	62.2	5.4	7091	4	US-09-085-273-40	Sequence 40, Appl
	38	62.2	5.4	7351	1	US-08-224-391-83	Sequence 83, Appl
	39	62.2	5.4	7351	1	US-08-484-304-83	Sequence 83, Appl
	40	62.2	5.4	7351	2	US-08-184-009-127	Sequence 127, Appl
	41	62.2	5.4	7351	2	US-08-566-398-39	Sequence 39, Appl
	42	62.2	5.4	7351	2	US-08-458-356-127	Sequence 127, Appl
	43	62.2	5.4	7351	2	US-08-658-665-39	Sequence 39, Appl
	44	62.2	5.4	7351	4	US-08-796-101-3	Sequence 3, Appl
	45	62.2	5.4	7351	4	US-08-460-736-127	Sequence 127, Appl

ALIGNMENTS

```
RESULT 1
US-08-991-789A-291/c
; Sequence 291, Application US/08991789A
; Patent No. 6225054
;  GENERAL INFORMATION:
;       APPLICANT: Frudakis, Tony N.
;                  Smith, John M.
;                  Reed, Steven G.
;       TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
;                           TREATMENT AND DIAGNOSIS OF BREAST CANCER
;       NUMBER OF SEQUENCES: 292
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Seed IP Law Group
;            STREET: 701 Fifth Avenue, Suite 6300
;            CITY: Seattle
;            STATE: Washington
;            COUNTRY: USA
;            ZIP: 98104-7092
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/08/991,789A
;            FILING DATE: 11-Dec-1997
;            CLASSIFICATION: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Potter, Jane E. R.
;            REGISTRATION NUMBER: 33,332
;            REFERENCE/DOCKET NUMBER: 210121.419C3
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (206) 622-4900
;            TELEFAX: (206) 682-6031
;    INFORMATION FOR SEQ ID NO: 291:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 1851 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       SEQUENCE DESCRIPTION: SEQ ID NO: 291:
US-08-991-789A-291

  Query Match            71.8%;  Score 829;  DB 4;  Length 1851;
  Best Local Similarity  95.1%;  Pred. No. 1.1e-233;
  Matches  878;  Conservative  0;  Mismatches  40;  Indels  5;  Gaps  2;

Qy     216 caggggagtggcaagagcaacgtgggcgcttctggagaccacaacaactctgctataaa 275
           ||||||||| ||||||||||||||||| ||||||||||||| | ||||  || ||||
```

```
Db     1081 CAGGGGAGCGGCAAGACCAACCTGCCCACTTCTGGAGACCACAACGACTCCTCTGTGAA 1022

Qy      276 gacgctcaggaacaagatgggcaagtggtgctgccactgcttccctgctgcaggggag 335
            ||| ||   ||| ||||| | |||||||||||||||||||||||||||||||| ||||||
Db     1021 GACGCTTGGAGCAAGAGGTGCAAGTGGTGCTGCCACTGCTTCCCCTGCTGCA-GGGGAG 963

Qy      336 cggcaagagcaaggtgggcgcttggggagactacgatgacagtgccttcatggagcccag 395
            ||||||||||| |||| ||||||||||||||||||||||||| ||||||||||| |||||
Db      962 CGGCAAGAGCAACGTGGTCGCTTGGGGAGACTACGATGACAGCGCCTTCATGGATCCCAG 903

Qy      396 gtaccacgtccgtggagaagatctggacaagctccacagagctgcctggtggggtaaagt 455
            ||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db      902 GTACCACGTCCATGGAGAAGATCTGGACAAGCTCCACAGAGCTGCCTGGTGGGGTAAAGT 843

Qy      456 ccccagaaaggatctcatcgtcatgctcagggacactgacgtgaacaagaaggacaagca 515
            |||||||||||||||||||||||||||||||||||| || |||||||||| |||||||||
Db      842 CCCCAGAAAGGATCTCATCGTCATGCTCAGGGACACGGATGTGAACAAGAGGGACAAGCA 783

Qy      516 aaagaggactgctctacatctggcctctgccaatgggaattcagaagtagtaaaactcct 575
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db      782 AAACAGGACTGCTCTACATCTGGCCTCTGCCAATGGGAATTCAGAAGTAGTAAAACTCGT 723

Qy      576 gctggacagacgatgtcaacttaatgtccttgacaacaaaaagaggacagctctgataaa 635
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db      722 GCTGGACAGACGATGTCAACTTAATGTCCTTGACAACAAAAAGAGGACAGCTCTGACAAA 663

Qy      636 ggccgtacaatgccaggaagatgaatgtgcgttaatgttgctggaacatggcactgatcc 695
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      662 GGCCGTACAATGCCAGGAAGATGAATGTGCGTTAATGTTGCTGGAACATGGCACTGATCC 603

Qy      696 aaatattccagatgagtatggaaataccactctgcactacgctatctataatgaagataa 755
            |||||||||||||||||||||||||||||||||| ||||| ||| |||| ||||||||||
Db      602 AAATATTCCAGATGAGTATGGAAATACCACTCTACACTATGCTGTCTACAATGAAGATAA 543

Qy      756 attaatggccaaagcactgctcttatatggtgctgatatcgaatcaaaaaacaagcatgg 815
            |||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db      542 ATTAATGGCCAAAGCACTGCTCTTATACGGTGCTGATATCGAATCAAAAAACAAGCATGG 483

Qy      816 cctcacaccactgttacttggtgtacatgagcaaaaacagcaagtcgtgaaattttaat 875
            ||||||||||||| |||||||| ||||||||||||||||||||||| |||||||||||||
Db      482 CCTCACACCACTGCTACTTGGTATACATGAGCAAAAACAGCAAGTGGTGAAATTTTTAAT 423

Qy      876 caagaaaaagcgaatttaaatgcactggatagatatggaaggactgctctcatacttgc 935
            ||||||||||||||||||||||| |||||||||||||||| ||||||||||||||||||
Db      422 CAAGAAAAAGCGAATTTAAATGCGCTGGATAGATATGGAAGAACTGCTCTCATACTTGT 363

Qy      936 tgtatgttgtggatcagcaagtatagtcagccttctacttgagcaaaatattgatgtatc 995
            |||||||||||||||||||| |||| ||||| ||||||||||||||||| ||||||||||
Db      362 TGTATGTTGTGGATCAGCAACTATACTCAGCCTCTACTTGAGCAAAATGTTGATGTATC 303

Qy      996 ttctcaagatctatctggacagacggccagagagtatgctgtttctagtcatcatcatgt 1055
            ||||||||    ||||||| ||||||||||||||||||||||||||||||||||||||||
Db      302 TTCTCAAC----ATCTGGAAACACGGCCAGACACTATCCTGTTTCTAGTCATCATCATGT 247

Qy     1056 aatttgccagttactttctgactacaaagaaaaacagatgctaaaaatctcttctgaaaa 1115
            ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db      246 AATTTGCCAGTTACTTTCTGACTACAAAGAAAAACAGATGTTAAAAATCTCTTCTGAAAA 187

Qy     1116 cagcaatccagaaaatgtctcaa 1138
            ||||||||||||| | | || ||
Db      186 CAGCAATCCAGAACAAGACTTAA 164
```

RESULT 2
US-08-991-789A-211/c
; Sequence 211, Application US/08991789A
; Patent No. 6225054
; GENERAL INFORMATION:
; APPLICANT: Frudakis, Tony N.
; Smith, John M.
; Reed, Steven G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TREATMENT AND DIAGNOSIS OF BREAST CANCER
; NUMBER OF SEQUENCES: 292
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed IP Law Group
; STREET: 701 Fifth Avenue, Suite 6300
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/991,789A
; FILING DATE: 11-Dec-1997
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Potter, Jane E. R.
; REGISTRATION NUMBER: 33,332
; REFERENCE/DOCKET NUMBER: 210121.419C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 211:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 454 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 211:
US-08-991-789A-211

```
  Query Match             36.2%;  Score 417.8;  DB 4;  Length 454;
  Best Local Similarity   96.2%;  Pred. No. 2.8e-113;
  Matches  428;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy      474 cgtcatgctcagggacactgacgtgaacaagaaggacaagcaaaagaggactgctctaca 533
            ||| ||||||||||||||||| |||||||||| ||||||||||||||||||||||||||
Db      451 CGTTATGCTCAGGGACACTGATGTGAACAAGAGGGACAAGCAAAAGAGGACTGCTCTACA 392

Qy      534 tctggcctctgccaatgggaattcagaagtagtaaaactcctgctggacagacgatgtca 593
            |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db      391 TCTGGCCTCTGCCAATGGGAATTCAGAAGTAGTAAAACTCGTGCTGGACAGACGATGTCA 332

Qy      594 acttaatgtccttgacaacaaaaagaggacagctctgataaaggccgtacaatgccagga 653
            |||||||||||||||||||||||||||||||||||| |||||||||||||||||||| .
Db      331 ACTTAATGTCCTTGACAACAAAAAGAGGACAGCTCTGACAAAGCCCGTACAATGCCAGCA 272

Qy      654 agatgaatgtgcgttaatgttgctggaacatggcactgatccaaatattccagatgagta 713
            ||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db      271 AGATCAATGTGCGTTAATGTTGCTGGAACATGGCACTGATCCAAATATTCCACATGAGTA 212

Qy      714 tggaaataccactctgcactacgctatctataatgaagataaattaatggccaaagcact 773
            |||| |||||||||| ||||| ||| |||| |||||||||||||||||||||||| |||||
Db      211 TGGAGATACCACTCTACACTATGCTGTCTACAATGAAGATAAATTAATGGCCAAAGCACT 152

Qy      774 gctcttatatggtgctgatatcgaatcaaaaaacaagcatggcctcacaccactgttact 833
            ||||||||| |||||||||||||||||||||||||||||||||||||||||||| ||||
Db      151 GCTCTTATACGGTGCTGATATCGAATCAAAAAACAAGCATGGCCTCACACCACTGCTACT 92

Qy      834 tggtgtacatgagcaaaaacagcaagtcgtgaaattttaatcaagaaaaagcgaattt 893
            |||| |||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db       91 TGGTATACATGAGCAAAAACAGCAAGTGGTGAAATTTTTAATCAAGAAAAAGCGAATTT 32

Qy      894 aaatgcactggatagatatggaagg 918
            | |||| |||||||||||| |||||
Db       31 AGATGCGCTGGATAGATATTGAAGG 7
```

RESULT 3
US-09-082-059-1

```
; Sequence 1, Application US/09082059A
; Patent No. 6225086
; GENERAL INFORMATION:
;  APPLICANT: Morrow, Jon S.
;  APPLICANT: Devarajan, Prasad
;  TITLE OF INVENTION: No. 6225086el Ankyrin Proteins and a Method for Their Identificat
;  FILE REFERENCE: 44574 5002 US
;  CURRENT APPLICATION NUMBER: US/09/082,059A
;  CURRENT FILING DATE: 1998 05 21
;  EARLIER APPLICATION NUMBER: 60/047356
;  EARLIER FILING DATE: 1997 05 21
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 3454
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (331)..(3394)
;   OTHER INFORMATION: DNA/protein sequence Human Kidney Ankyrin G119
US 09 082 059 1

  Query Match             7.3%;  Score 84.6;  DB 4;  Length 3454;
  Best Local Similarity   49.9%;  Pred. No. 5.9e 15;
  Matches  213;  Conservative    0;  Mismatches  214;  Indels    0;  Gaps    0;

Qy     511 aaacaaaagaggactgctctacatctggcctctgccaatgggaattcagaagtagtaaaa 570
Db     680 [illegible] 739

Qy     571 ctcctgctggacaagcgatgtcaacttaatgtccttgacaacaaaaagaggacagctctg 630
Db     740 [illegible] 799

Qy     631 ataaaggccgtacaatgccaggaagatgaatgtgcgttaatgttgctggaacatggcact 690
Db     800 [illegible] 859

Qy     691 gatccaaatattccagatgagtatggaaataccactctgcactacgctatctataatgaa 750
Db     860 [illegible] 919

Qy     751 gataaattaatggccaaagcactgctcttatatggtgctgatatcgaatcaaaaaacaag 810
Db     920 [illegible] 979

Qy     811 catgggctcacaccactgttacttggtgtacatgagcaaaaacagcaagtcgtgaaattt 870
Db     980 [illegible] 1039

Qy     871 ttaatcaagaaaaaagcgaatttaaatgcactggatagatatggaaggactgctctcata 930
Db    1040 [illegible] 1099

Qy     931 cttgctg 937
Db    1100 ttggctg 1106

RESULT 4
US 09 065 474 138
; Sequence 138, Application US/09065474
; Patent No. 6263509
; GENERAL INFORMATION:
;   APPLICANT: Tang, Liang
;   APPLICANT: Bleham, E. Scot
;   TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
;   TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
;   TITLE OF INVENTION: USES THEREOF
;   NUMBER OF SEQUENCES: 171
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Carol Talkington Verser, Ph.D.
;     ADDRESSEE: Heska Corporation
;     STREET: 1825 Sharp Point Drive
;     CITY: Fort Collins
;     STATE: Colorado
;     COUNTRY: USA
;     ZIP: 80525
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: Windows 95
;     SOFTWARE: WordPerfect for Windows, Version 7.0
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/09/065,474
;     FILING DATE: 24 APR 1998
;     CLASSIFICATION:
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Verser, Carol Talkington
;     REGISTRATION NUMBER: 37,459
;     REFERENCE/DOCKET NUMBER: HW 5 C1
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 970/493 7272
;     TELEFAX: 970/484 9505
; INFORMATION FOR SEQ ID NO: 138:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 1056 nucleotides
;     TYPE: nucleic acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;   MOLECULE TYPE: cDNA
;   FEATURE:
;     NAME/KEY: CDS
;     LOCATION: 1..1056
US 09 065 474 138

  Query Match             7.2%;  Score 82.6;  DB 3;  Length 1056;
  Best Local Similarity   47.6%;  Pred. No. 1.2e 14;
  Matches  244;  Conservative    0;  Mismatches  269;  Indels    0;  Gaps    0;

Qy     523 actgctctacatctggcctctgccaatgggaattcagaagtagtaaaactcctgctggac 582
Db     334 ACGGCGTTGCATTTAGCTGCTAAGTATGGCAATTTGCTGGTGGAAATCATTGCTACAA 393

Qy     583 aagcgatgtcaacttaatgtccttgacaacaaaaagaggacagctctgataaaggccgta 642
Db     394 CGAGGAACACCGGTTGACATTGAAGGAAGAATAGGTAACAC[illegible]AGTAGCGGCA 453

Qy     643 caatgccaggaagatgaatgtgcgttaatgttgctggaacatggcactgatccaaatatt 702
Db     454 CATTACAATAACGACAAGGTAG[illegible] 513

Qy     703 ccagatgagtatggaaataccactctgcactacgctatctataatgaagataaattaatg 762
Db     514 GCTGCTAAGAATGGGTACACTCCTTTACATATTGCCGGGAAGAAGAATAGATGGATATT 573

Qy     763 gccaaagcactgctcttatatggtgctgatatcgaatcaaaaaacaagcatgggctcaca 822
Db     574 GCTAGCACTGTCCTTCATTATAAGGCAAATGCGAATGCTGAAAGCAAAGCTGGCTTACA 633

Qy     823 ccactgttacttggtgtacatgagcaaaaacagcaagtcgtgaaatttttaatcaagaaa 882
Db     634 CCACTTCATCTTGCCGCCCAGGAGGGCCATCGCGAAATGGCTGCGTTATTAATTGAAAAT 693

Qy     883 aaagcgaatttaaatgcactggatagatatggaaggactgctctcatacttgctgtatat 942
Db     694 GGAGCAAAAGTTGGAGCTCAGGCAAGGAATGGCTTGACACCAATGCATTTATGTGCACAG 753

Qy     943 tgtggatcagcaagtatagtcagccttctacttgagcaaaatattgatatatctctgcaa 1002
Db     754 GAGGATCGTGTGAGCGTAGCAGAAGAACTAGTGAAAGAAAA[illegible] 813

Qy    1003 gatctatctggacagacggccaaagagtatgct 1035
```

```
Db     814 ACGAAAGCAGGATATACGCCGTTACATGTTGCT 846

RESULT   5
US-09-065-474-140/c
; Sequence 140, Application US/09065474
; Patent No. 6063599
;  GENERAL INFORMATION:
;    APPLICANT:  Tang, Liang
;    APPLICANT:  Blehm, E. Scot
;    TITLE OF INVENTION:  DIROFILARIA AND BRUGIA ANKYRIN
;    TITLE OF INVENTION:  PROTEINS, NUCLEIC ACID MOLECULES, AND
;    TITLE OF INVENTION:  USES THEREOF
;    NUMBER OF SEQUENCES:  171
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Carol Talkington Verser, Ph.D.
;      ADDRESSEE:  Heska Corporation
;      STREET:  1825 Sharp Point Drive
;      CITY:  Fort Collins
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80525
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect for Windows, Version 7.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/065,474
;      FILING DATE:  24-APR-1998
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Verser, Carol Talkington
;      REGISTRATION NUMBER:  37,459
;      REFERENCE/DOCKET NUMBER:  HW-5-C1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  970/493-7272
;      TELEFAX:  970/484-9505
;  INFORMATION FOR SEQ ID NO:  140:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1056 nucleotides
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-065-474-140

  Query Match             7.2%;  Score 82.6;  DB 3;  Length 1056;
  Best Local Similarity   47.6%;  Pred. No. 1.2e-14;
  Matches  244;  Conservative    0;  Mismatches  269;  Indels    0;  Gaps    0;

Qy     523 actgctctacatctggcctctgccaatgggaattcaaagtagtaaaactcctgctggac 582
Db     723 ACGCCGTTGCATTTAGCTGCTAAGTATGGCAATTTGCCGGTCGCGAAATCATTGCTAGAA 664

Qy     583 agacgatgtcaacttaatgtccttgacaacaaaaagaggacagctctgataaaggccgta 642
Db     663 CGAGGAACACCGGTTGACATTGAAGGCAAGAATCAGGTAACACCTCTGCATGTAGCGGCA 604

Qy     643 caatgccaggaagatgaatgtgcgttaatgttgctggaacatggcactgatccaaatatt 702
Db     603 CATTACAATAACGACAAGGTAGCATTGTTACTTCTAGAAAATGGTGCTTCTGCACATGCC 544

Qy     703 ccagatgagtatggaaataccactctgcactacgctatctataatgaagataaattaatg 762
Db     543 GCTGCCAACAATCGCTACACTCCTTTACATATTCCCCCAACAACAATCACATGCATATT 484

Qy     763 gccaagcactgctcttatatgtgctgatatcgaatcaaaaacagcatggcctcaca 822
Db     483 GCTAGCACTCTCCTTCATTATAAGCCAAATGCGAATGCTGAAAGCAAAGCTGGCTTTACA 424

Qy     823 ccactgttacttggtgtacatgagcaaaaacagcaagtcgtgaaattttaatcaagaaa 882
Db     423 CCACTTCATCTTGCCGCCCAGGAGGGCCATCGCGAAATGGCTGCGTTATTAATTGAAAAT 364

Qy     883 aaagcgaatttaaatgcactggatagatatggaaggactgctctcatacttgctgtatgt 942
Db     363 GCAGCAAAAGTTGCACCTCAGCCAAGCAATGGCTTGACACCAATCCATTTATTGGCACAG 304

Qy     943 tgtggatcagcaagtatagtcagccttctacttgagcaaaatattgatgtatcttctcaa 1002
Db     303 GAGGATCGTGTGAGCGTAGCAGAAGAACTAGTGAAAGAAAACGCAGCCATTGATCCCAAA 244

Qy    1003 gatctatctggacagacggccagagagtatgct 1035
Db     243 ACGAAAGCAGGATATACGCCGTTACATGTTGCT 211

RESULT   6
US-09-031-485-35
; Sequence 35, Application US/09031485
; Patent No. 5824306
;  GENERAL INFORMATION:
;    APPLICANT:  Tang, Liang
;    APPLICANT:  Blehm, E. Scot
;    TITLE OF INVENTION:  DIROFILARIA AND BRUGIA ANKYRIN
;    TITLE OF INVENTION:  PROTEINS, NUCLEIC ACID MOLECULES, AND
;    TITLE OF INVENTION:  USES THEREOF
;    NUMBER OF SEQUENCES:  85
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Carol Talkington Verser, Ph.D.
;      ADDRESSEE:  Heska Corporation
;      STREET:  1825 Sharp Point Drive
;      CITY:  Fort Collins
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80525
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect for Windows, Version 7.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/031,485
;      FILING DATE:
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/847,429
;      FILING DATE:  24-APR-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Verser, Carol Talkington
;      REGISTRATION NUMBER:  37,459
;      REFERENCE/DOCKET NUMBER:  HW-5
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  970/493-7272
;      TELEFAX:  970/484-9505
;  INFORMATION FOR SEQ ID NO:  35:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5235 nucleotides
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-031-485-35

  Query Match             7.2%;  Score 82.6;  DB 1;  Length 5235;
  Best Local Similarity   47.6%;  Pred. No. 2.9e-14;
  Matches  244;  Conservative    0;  Mismatches  269;  Indels    0;  Gaps    0;

Qy     523 actgctctacatctggcctctgccaatgggaattcaaagtagtaaaactcctgctggac 582
Db    1639 ACGCCGTTGCATTTAGCTGCTAAGTATGGCAATTTGCCGGTCGCGAAATCATTGCTAGAA 1698
```

```
Qy    583 agacgatgtcaacttaatgtccttgacaacaaaaagaggacagctctgataaaggccgta 642
Db   1699 CGAGGAACACCGGTTGACATTGAAGGAAGAATCAGGTAACACCTCTGCATGTAGCGGCA 1758

Qy    643 caatgccaggaagatgaatgtgcgttaatgttgctggaacatggcactgatccaaatatt 702
Db   1759 CATTACAATAACGACAAGGTAGCATTGTTACTTCTAGAAAATGGTGCTTCTGCACATGCC 1818

Qy    703 ccagatgagtatggaaataccactctgcactacgctatctataatgaagataaattaatg 762
Db   1819 GCTGCCAAGAATGGGTACACTCCTTTACATATTGCCGCGAAGAAGAATCAGATGGATATT 1878

Qy    763 gccaaagcactgctcttatatggtgctgatatcgaatcaaaaaacaagcatggcctcaca 822
Db   1879 GCTAGCACTCTCCTTCATTATAAGGCAAATGCGAATGCTGAAAGCAAAGCTGGCTTTACA 1938

Qy    823 ccactgttacttggtgtacatgagcaaaaacagcaagtcgtgaaatttttaatcaagaaa 882
Db   1939 CCACTTCATCTTGCCGCCCAGGAGGGCCATCGCGAAATGGCTGCGTTATTAATTGAAAAT 1998

Qy    883 aaagcgaatttaaatgcactggatagatatggaaggactgctctcatacttgctgtatgt 942
Db   1999 GGAGCAAAAGTTGGAGCTCAGGCAAGGAATGGCTTGACACCAATGCATTTATGTGCACAG 2058

Qy    943 tgtggatcagcaagtatagtcagccttctacttgagcaaaatattgatgtatcttctcaa 1002
Db   2059 GAGGATCGTGTGAGCGTAGCAGAAGAACTAGTGAAAGAAAACGCAGCCATTGATCCCAAA 2118

Qy   1003 gatctatctggacagacggccagagagtatgct 1035
Db   2119 ACGAAAGCAGGATATACGCCGTTACATGTTGCT 2151
```

```
RESULT 7
US-09-031-485-36/c
; Sequence 36, Application US/09031485
; Patent No. 5824306
; GENERAL INFORMATION:
;    APPLICANT: Tang, Liang
;    APPLICANT: Blehm, E. Scot
;    TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
;    TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
;    TITLE OF INVENTION: USES THEREOF
;    NUMBER OF SEQUENCES: 85
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE: Carol Talkington Verser, Ph.D.
;      ADDRESSEE: Heska Corporation
;      STREET: 1825 Sharp Point Drive
;      CITY: Fort Collins
;      STATE: Colorado
;      COUNTRY: USA
;      ZIP: 80525
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE: Floppy disk
;      COMPUTER: IBM PC compatible
;      OPERATING SYSTEM: Windows 95
;      SOFTWARE: WordPerfect for Windows, Version 7.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER: US/09/031,485
;      FILING DATE:
;      CLASSIFICATION: 530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: 08/847,429
;      FILING DATE: 24 APR 1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME: Verser, Carol Talkington
;      REGISTRATION NUMBER: 37,459
;      REFERENCE/DOCKET NUMBER: HW 5
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE: 970/493 7272
;      TELEFAX: 970/484 9505
;  INFORMATION FOR SEQ ID NO: 36:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH: 5235 nucleotides
;      TYPE: nucleic acid
;      STRANDEDNESS: single
;      TOPOLOGY: linear
;    MOLECULE TYPE: cDNA
US-09-031-485-36

  Query Match             7.2%;  Score 82.6;  DB 1;  Length 5235;
  Best Local Similarity   47.6%;  Pred. No. 2.9e-14;
  Matches  244;  Conservative    0;  Mismatches  269;  Indels    0;  Gaps    0;

Qy    523 actgctctacatctggcctctgccaatgggaattcagaagtagtaaaactcctgctggac 582
Db   3597 ACGCCGTTGCATTTAGCTGCTAAGTATGGCAATTTGCCGGTGCGAAATCATTGGTAGAA 3538

Qy    583 agacgatgtcaacttaatgtccttgacaacaaaaagaggacagctctgataaaggccgta 642
Db   3537 CGAGGAACACCGGTTGACATTGAAGGAAGAATCAGGTAACACCTCTGCATGTAGCGGCA 3478

Qy    643 caatgccaggaagatgaatgtgcgttaatgttgctggaacatggcactgatccaaatatt 702
Db   3477 CATTACAATAACGACAAGGTAGCATTGTTACTTCTAGAAAATGGTGCTTCTGCACATGCC 3418

Qy    703 ccagatgagtatggaaataccactctgcactacgctatctataatgaagataaattaatg 762
Db   3417 GCTGCCAAGAATGGGTACACTCCTTTACATATTGCCGCGAAGAAGAATCAGATGGATATT 3358

Qy    763 gccaaagcactgctcttatatggtgctgatatcgaatcaaaaaacaagcatggcctcaca 822
Db   3357 GCTAGCACTCTCCTTCATTATAAGGCAAATGCGAATGCTGAAAGCAAAGCTGGCTTTACA 3298

Qy    823 ccactgttacttggtgtacatgagcaaaaacagcaagtcgtgaaatttttaatcaagaaa 882
Db   3297 CCACTTCATCTTGCCGCCCAGGAGGGCCATCGCGAAATGGCTGCGTTATTAATTGAAAAT 3238

Qy    883 aaagcgaatttaaatgcactggatagatatggaaggactgctctcatacttgctgtatgt 942
Db   3237 GGAGCAAAAGTTGGAGCTCAGGCAAGGAATGGCTTGACACCAATGCATTTATGTGCACAG 3178

Qy    943 tgtggatcagcaagtatagtcagccttctacttgagcaaaatattgatgtatcttctcaa 1002
Db   3177 GAGGATCGTGTGAGCGTAGCAGAAGAACTAGTGAAAGAAAACGCAGCCATTGATCCCAAA 3118

Qy   1003 gatctatctggacagacggccagagagtatgct 1035
Db   3117 ACGAAAGCAGGATATACGCCGTTACATGTTGCT 3085

RESULT 8
US-08-847-429A-35
; Sequence 35, Application US/08847429A
; Patent No. 5827692
; GENERAL INFORMATION:
;    APPLICANT: Tang, Liang
;    APPLICANT: Blehm, E. Scot
;    TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
;    TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
;    TITLE OF INVENTION: USES THEREOF
;    NUMBER OF SEQUENCES: 85
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE: Carol Talkington Verser, Ph.D.
;      ADDRESSEE: Heska Corporation
;      STREET: 1825 Sharp Point Drive
;      CITY: Fort Collins
;      STATE: Colorado
;      COUNTRY: USA
;      ZIP: 80525
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE: Floppy disk
;      COMPUTER: IBM PC compatible
;      OPERATING SYSTEM: Windows 95
;      SOFTWARE: WordPerfect for Windows, Version 7.0
```

```
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/847,429A
;      FILING DATE:  24-APR-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Verser, Carol Talkington
;      REGISTRATION NUMBER:  37,459
;      REFERENCE/DOCKET NUMBER:  HW-5
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  970/493-7272
;      TELEFAX:  970/484-9505
;  INFORMATION FOR SEQ ID NO:  35:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5235 nucleotides
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-847-429A-35

  Query Match             7.2%;  Score 82.6;  DB 1;  Length 5235;
  Best Local Similarity   47.6%;  Pred. No. 2.9e-14;
  Matches  244;  Conservative    0;  Mismatches  269;  Indels    0;  Gaps    0;

Qy     523 actgctctacatctqqcctctqccaatqqqaattcaqaaqtaqtaaaactcctgctggac 582
           ||  |  | ||| | ||  ||    |||| ||||     || |  |||    |||| ||
Db    1639 ACGCCGTTGCATTTAGCTGCTAAGTATGGCAATTTGCCGGTCGCGAAATCATTGCTAGAA 1698

Qy     583 aqacgatgtcaacttaatgtccttgacaacaaaaaqaqgacaqctctgataaaqqccgta 642
           || ||   |   || | |    | ||| ||  ||   ||| ||||      || | |
Db    1699 CGAGGAACACCGGTTGACATTGAAGGCAAGAATCAGGTAACACCTCTGCATCTAGCGCCA 1758

Qy     643 caatgccaqgaaqatqaatqtqcqttaatgttgctqqaacatqqcactqatccaaatatt 702
           || | | |  | ||  |   || ||  |  | || ||| |||| ||  | | ||
Db    1759 CATTACAATAACGACAAGGTAGCATTGTTACTTCTAGAAAATGGTGCTTCTGCACATGCC 1818

Qy     703 ccaqatqaqtatqqaaatacccactctqcactacgctatctataatqaaqataaattaatg 762
           | |   || |||| | ||  || | ||    ||     | ||  |  | |     ||
Db    1819 GCTGCCAAGAATGGGTACACTCCTTTACATATTGCCGCGAAGAAGAATCAGATGGATATT 1878

Qy     763 qccaaaqcactgctcttatatqgtgctgatatcqaatcaaaaaacaaqcatqgcctcaca 822
           || |   | || ||    |||  ||  ||    | |  ||| |||   |||| | |||
Db    1879 GCTAGCACTCTCCTTCATTATAAGGCAAATGCGAATGCTGAAAGCAAAGCTGGCTTTACA 1938

Qy     823 ccactgttacttqgtgtacatqaqcaaaaacaqcaagtcgtgaaatttttaatcaagaaa 882
           |||||    ||||  |  || |||    | |   || | |     || |||||  | ||
Db    1939 CCACTTCATCTTGCCGCCCAGGAGGGCCATCGCGAAATGGCTGCGTTATTAATTGAAAAT 1998

Qy     883 aaaqcqaatttaaatqcactqaataqatatqqaaqgactqctctcatacttgctgtatgt 942
           ||| ||  |   || | ||  ||  |||| |||  | |   |   | ||
Db    1999 GGAGCAAAAGTTGGAGCTCAGGCAAGGAATGGCTTGACACCAATGCATTTATGTGCACAG 2058

Qy     943 tqtqqatcaqcaaqtataqtcaqccttctacttqaqcaaaaatattqatqtatcttctcaa 1002
             |     |  || |||        ||| |  |  ||||    |   |  | |  ||
Db    2059 GAGGATCGTGTGACGCTAGCAGAAGAACTAGTGAAAGAAAACGCAGCCATTGATCCCAAA 2118

Qy    1003 qatctatctqqacaqacqqccaqaqaqtatqct 1035
              | | ||| | ||| |   | |   ||||
Db    2119 ACGAAAGCAGGATATACGCCGTTACATGTTGCT 2151


RESULT   9
US-08-847-429A-36/c
; Sequence 36, Application US/08847429A
; Patent No. 5827692
;  GENERAL INFORMATION:
;    APPLICANT:  Tang, Liang
;    APPLICANT:  Blehm, E. Scot
;    TITLE OF INVENTION:  DIROFILARIA AND BRUGIA ANKYRIN
;    TITLE OF INVENTION:  PROTEINS, NUCLEIC ACID MOLECULES, AND
;    TITLE OF INVENTION:  USES THEREOF
;    NUMBER OF SEQUENCES:  85
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Carol Talkington Verser, Ph.D.
;      ADDRESSEE:  Heska Corporation
;      STREET:  1825 Sharp Point Drive
;      CITY:  Fort Collins
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80525
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect for Windows, Version 7.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/847,429A
;      FILING DATE:  24-APR-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Verser, Carol Talkington
;      REGISTRATION NUMBER:  37,459
;      REFERENCE/DOCKET NUMBER:  HW-5
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  970/493-7272
;      TELEFAX:  970/484-9505
;  INFORMATION FOR SEQ ID NO:  36:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5235 nucleotides
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-847-429A-36

  Query Match             7.2%;  Score 82.6;  DB 1;  Length 5235;
  Best Local Similarity   47.6%;  Pred. No. 2.9e-14;
  Matches  244;  Conservative    0;  Mismatches  269;  Indels    0;  Gaps    0;

Qy     523 actgctctacatctqqcctctqccaatqqqaattcaqaaqtaqtaaaactcctgctggac 582
           ||  |  | ||| | ||  ||    |||| ||||     || |  |||    |||| ||
Db    3597 ACGCCGTTGCATTTAGCTGCTAAGTATGGCAATTTGCCGGTCGCGAAATCATTGCTAGAA 3538

Qy     583 aqacgatgtcaacttaatgtccttgacaacaaaaaqaqgacaqctctgataaaqqccgta 642
           || ||   |   || | |    | ||| ||  ||   ||| ||||      || | |
Db    3537 CGAGGAACACCGGTTGACATTGAAGGCAAGAATCAGGTAACACCTCTGCATGTAGCGGCA 3478

Qy     643 caatgccaqgaaqatqaatqtqcqttaatgttgctqqaacatqqcactqatccaaatatt 702
           || | | |  | ||  |   || ||  |  | || ||| |||| ||  | | ||
Db    3477 CATTACAATAACGACAAGGTAGCATTGTTACTTCTAGAAAATGGTGCTTCTGCACATGCC 3418

Qy     703 ccaqatqaqtatqqaaatacccactctqcactacgctatctataatqaaqataaattaatg 762
           | |   || |||| | ||  || | ||    ||     | ||  |  | |     ||
Db    3417 GCTGCCAAGAATGGGTACACTCCTTTACATATTGCCGCGAAGAAGAATCAGATGGATATT 3358

Qy     763 qccaaaqcactgctcttatatqgtgctgatatcqaatcaaaaaacaaqcatqgcctcaca 822
           || |   | || ||    |||  ||  ||    | |  ||| |||   |||| | |||
Db    3357 GCTAGCACTCTCCTTCATTATAAGGCAAATGCGAATGCTGAAAGCAAAGCTGGCTTTACA 3298

Qy     823 ccactgttacttqgtgtacatqaqcaaaaacaqcaagtcgtgaaatttttaatcaagaaa 882
           |||||    ||||  |  || |||    | |   || | |     || |||||  | ||
Db    3297 CCACTTCATCTTGCCGCCCAGGAGGGCCATCGCGAAATGGCTGCGTTATTAATTGAAAAT 3238

Qy     883 aaaqcqaatttaaatqcactqaatagatatqqaaqgactgctctcatacttgctgtatgt 942
           ||| ||  |   || | ||  ||  |||| |||  | |   |   | ||
Db    3237 GGAGCAAAAGTTGGAGCTCAGGCAAGGAATGGCTTGACACCAATGCATTTATGTGCACAG 3178

Qy     943 tqtqqatcaqcaaqtataatcaqccttctacttqaqcaaaaatattqatqtatcttctcaa 1002
             |     |  || |||        ||| |  |  ||||    |   |  | |  ||
Db    3177 GAGGATCGTGTGACGCTAGCAGAAGAACTAGTGAAAGAAAACGCAGCCATTGATCCCAAA 3118

Qy    1003 qatctatctqqacaqacqqccaqaqaqtatqct 1035
              | | ||| | ||| |   | |   ||||
Db    3117 ACGAAAGCAGGATATACGCCGTTACATGTTGCT 3085
```

```
RESULT 10
US-09-065-474-35
; Sequence 35, Application US/09065474
; Patent No. 6063599
;  GENERAL INFORMATION:
;    APPLICANT:  Tang, Liang
;    APPLICANT:  Blehm, E. Scot
;    TITLE OF INVENTION:  DIROFILARIA AND BRUGIA ANKYRIN
;    TITLE OF INVENTION:  PROTEINS, NUCLEIC ACID MOLECULES, AND
;    TITLE OF INVENTION:  USES THEREOF
;    NUMBER OF SEQUENCES:  171
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Carol Talkington Verser, Ph.D.
;      ADDRESSEE:  Heska Corporation
;      STREET:  1825 Sharp Point Drive
;      CITY:  Fort Collins
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80525
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect for Windows, Version 7.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/065,474
;      FILING DATE:  24 APR 1998
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Verser, Carol Talkington
;      REGISTRATION NUMBER:  37,459
;      REFERENCE/DOCKET NUMBER:  HW 5 C1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  970/493 7272
;      TELEFAX:  970/484 9505
;  INFORMATION FOR SEQ ID NO:  35:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5235 nucleotides
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-065-474-35

  Query Match             7.2%;  Score 82.6;  DB 3;  Length 5235;
  Best Local Similarity   47.6%;  Pred. No. 2.9e-14;
  Matches  244;  Conservative    0;  Mismatches  269;  Indels    0;  Gaps    0;

Qy        523 actgctctacatctggcctctgccaatgggaattcagaagtagtaaaactcctgctggac 582
Db       1639 ACGCCGTTGCATTTAGCTGCTAAGTATGGCAATTTGCCGGTCGCGAAATCATTGCTAGAA 1698

Qy        583 agacgatatgtcaacttaatgtccttgacaacaaaaagaggacagctctgataaaggccgta 642
Db       1699 CGAGGAACACCGGTTGACATTGAAGGCAAGAATCAGGTAACACCTCTGCATGTAGCGGCA 1758

Qy        643 caatgccaggaagatgaatgtgcgttaatattgctggaacatggcactgatccaaatatt 702
Db       1759 CATTACTAATAACGACAAGGTAGCATTGTTACTTCTAGAAAATGGTGCTTCTGCACATGCC 1818

Qy        703 ccagatgagtatggaaatacccactctgcactacgctatctataatgaagataaattaatg 762
Db       1819 GCTGGCAAGAATGGGTACACTCCTTTACATATTGCCGCGAAGAAGAATCAGATGGATATT 1878

Qy        763 gccaaagcactgctcttatatggtgctgatatcgaatcaaaaaacaagcatggcctcaca 822
Db       1879 GCTAGCATGCTGCTTCATTATAAGGCAAATGCGAATGCTGAAAGCAAAGCTGGCTTTACA 1938

Qy        823 ccactgttacttgatgtacatgagcaaaacagcaagtcgtgaaatttttaatcaagaaa 882
Db       1939 CCACTTCATCTTGCCGCCCAGGAGGGCCATCGGGAAATGGCTGCGTTATTAATTGAAAAT 1998

Qy        883 aaagcgaatttaaatgcactggatagatatggaaggactgctctcatacttgctgtatat 942
Db       1999 GGAGCAAAAGTTGGAGCTCAGGCAAGGAATGGCTTGACACCAATGCATTTATGTGCACAG 2058

Qy        943 tgtggatcagcaagtatagtcagccttctacttgagcaaaatattgatattatcttctcaa 1002
Db       2059 GAGGATCGGGTGAGCGTAGCAGAAGAACTAGTGAAAGAAAACGCAGCCATTGATCCCAAA 2118

Qy       1003 gatctatctggacagacggccagagagtatgct 1035
Db       2119 ACGAAAGCAGGATATACGCCGTTACATGTTGCT 2151


RESULT 11
US-09-065-474-36/c
; Sequence 36, Application US/09065474
; Patent No. 6063599
;  GENERAL INFORMATION:
;    APPLICANT:  Tang, Liang
;    APPLICANT:  Blehm, E. Scot
;    TITLE OF INVENTION:  DIROFILARIA AND BRUGIA ANKYRIN
;    TITLE OF INVENTION:  PROTEINS, NUCLEIC ACID MOLECULES, AND
;    TITLE OF INVENTION:  USES THEREOF
;    NUMBER OF SEQUENCES:  171
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Carol Talkington Verser, Ph.D.
;      ADDRESSEE:  Heska Corporation
;      STREET:  1825 Sharp Point Drive
;      CITY:  Fort Collins
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80525
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect for Windows, Version 7.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/065,474
;      FILING DATE:  24 APR 1998
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Verser, Carol Talkington
;      REGISTRATION NUMBER:  37,459
;      REFERENCE/DOCKET NUMBER:  HW 5 C1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  970/493 7272
;      TELEFAX:  970/484 9505
;  INFORMATION FOR SEQ ID NO:  36:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5235 nucleotides
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-065-474-36

  Query Match             7.2%;  Score 82.6;  DB 3;  Length 5235;
  Best Local Similarity   47.6%;  Pred. No. 2.9e-14;
  Matches  244;  Conservative    0;  Mismatches  269;  Indels    0;  Gaps    0;

Qy        523 actgctctacatctggcctctgccaatgggaattcagaagtagtaaaactcctgctggac 582
Db       3597 ACGCCGTTGCATTTAGCTGCTAAGTATGGCAATTTGCCGGTCGCGAAATCATTGCTAGAA 3538

Qy        583 agacgatatgtcaacttaatgtccttgacaacaaaaagaggacagctctgataaaggccgta 642
Db       3537 CGAGGAACACCGGTTGACATTGAAGGCAAGAATCAGGTAACACCTCTGCATGTAGCGGCA 3478

Qy        643 caatgccaggaagatgaatgtgcgttaatattgctggaacatggcactgatccaaatatt 702
```

```
Db    3477 CATTACAATAACGACAAGGTAGCATTGTTACTTCTAGAAAAIGGIGCIICIGCACATGCC 3418

Qy     703 ccaqatqaqtatatqqaaaataccactctqcactacqctatctatatqaaqaataaattaatq 762
Db    3417 GCTGCCAAGAATGGGTACACTCCTTTACATATTGCCGCGAAGAAGAATCAGATGGATATT 3358

Qy     763 qccaaaqcactqctctttatatqqtqctqatatcqaaatcaaaaaaacaaqcaatqqcctcaca 822
Db    3357 GCTAGCACTCTGCTTCATTATAAGGCAAATGCGAATGCTGAAAGCAAAGCTGGCTTTACA 3298

Qy     823 ccactgttacttgtgtacatgaqcaaaaacaqcaaqtcgtqaaattttaatcaaqaaa 882
Db    3297 CCACTTCATCTTGCCGCCCAGGAGGGCCATCGCGAAATGGCTGCGTTATTAATTGAAAAT 3238

Qy     883 aaaqcqaatttaaatqcactqqataqatatqqaaqgactqctctcatacttqctqtatqt 942
Db    3237 GGAGCAAAAGTTGCAGCTCAGGCAAGGAATGGCTTGACACCAATGCATTTATGTGCACAG 3178

Qy     943 tqtqqatcaqcaaqtatagtcagccttctacttgagcaaaatattgatqtatctttctcaa 1002
Db    3177 GAGGATCGTGTGAGCGTAGCAGAAGAACTAGTGAAAGAAAACGCAGCCATTGATCCCAAA 3118

Qy    1003 qatctatctqqacaqacqqccaqaqaqtatqct 1035
Db    3117 ACGAAAGCAGGATATACGCCGTTACATGTTGCT 3085


RESULT 12
US-09-031-485-32
; Sequence 32, Application US/09031485
; Patent No. 5824306
;  GENERAL INFORMATION:
;    APPLICANT:  Tang, Liang
;    APPLICANT:  Blehm, E. Scot
;    TITLE OF INVENTION:  DIROFILARIA AND BRUGIA ANKYRIN
;    TITLE OF INVENTION:  PROTEINS, NUCLEIC ACID MOLECULES, AND
;    TITLE OF INVENTION:  USES THEREOF
;    NUMBER OF SEQUENCES:  85
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Carol Talkington Verser, Ph.D.
;      ADDRESSEE:  Heska Corporation
;      STREET:  1825 Sharp Point Drive
;      CITY:  Fort Collins
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80525
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect for Windows, Version 7.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/031,485
;      FILING DATE:
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/847,429
;      FILING DATE:  24-APR-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Verser, Carol Talkington
;      REGISTRATION NUMBER:  37,459
;      REFERENCE/DOCKET NUMBER:  HW-5
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  970/493-7272
;      TELEFAX:  970/484-9505
;  INFORMATION FOR SEQ ID NO:  32:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5503 nucleotides
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  51..5285
US-09-031-485-32


  Query Match             7.2%;  Score 82.6;  DB 1;  Length 5503;
  Best Local Similarity   47.6%;  Pred. No. 2.9e-14;
  Matches  244;  Conservative    0;  Mismatches  269;  Indels    0;  Gaps    0;

Qy     523 actgctctacatctqqcctctqccaatqqqaattcaaaqtaqtaaaactcctqctqtac 582
Db    1689 ACGCCGTTGCATTTAGCTGCTAAGTATGGCAATTTGCCGGTCGCGAAATCATTGCTAGAA 1748

Qy     583 agacqatqtqaaacttaatgtccttgacaacaaaaaqaqqacaqctctgataaqqccata 642
Db    1749 CGAGCAACACCGGTTGACATTGAAGGCAAGAATCAGTAACACCTCTGCATGTAGCGGCA 1808

Qy     643 caatqccaqqaaqatqaatatqtqcqttaatqttqctqqaaacatqqcactqatccaaatatt 702
Db    1809 CATTACAATAACGACAAGGTAGCATTGTTACTTCTAGAAAATGGTGCTTCTGCACATGCC 1868

Qy     703 ccaqatqaqtatqqaaatacccactctqcactacqctatctatatqaaqaataaattaatq 762
Db    1869 GCTGCCAAGAATGGGTACACTCCTTTACATATTGCCGCGAAGAAGAATCAGATGGATATT 1928

Qy     763 qccaaaqcactqctctttatatqqtqctqatatcqaatcaaaaaacaaqcaatqqcctcaca 822
Db    1929 GCTAGCACTCTGCTTCATTATAAGGCAAATGCGAATGCTGAAAGCAAAGCTGGCTTTACA 1988

Qy     823 ccactgttacttgtgtacatgaqcaaaaacaqcaaqtcgtqaaattttaatcaaqaaa 882
Db    1989 CCACTTCATCTTGCCGCCCAGGAGGGCCATCGCGAAATGGCTGCGTTATTAATTGAAAAT 2048

Qy     883 aaaqcqaatttaaatqcactqqataqatatqqaaqgactqctctcatacttqctqtatqt 942
Db    2049 GGAGCAAAAGTTGCAGCTCAGGCAAGGAATGGCTTGACACCAATGCATTTATGTGCACAG 2108

Qy     943 tqtqqatcaqcaaqtatagtcagccttctacttgagcaaaatattgatqtatctttctcaa 1002
Db    2109 GAGGATCGTGTGAGCGTAGCAGAAGAACTAGTGAAAGAAAACGCAGCCATTGATCCCAAA 2168

Qy    1003 qatctatctqqacaqacqqccaqaqaqtatqct 1035
Db    2169 ACGAAAGCAGGATATACGCCGTTACATGTTGCT 2201


RESULT 13
US-09-041-485-34/c
; Sequence 34, Application US/09031485
; Patent No. 5824306
;  GENERAL INFORMATION:
;    APPLICANT:  Tang, Liang
;    APPLICANT:  Blehm, E. Scot
;    TITLE OF INVENTION:  DIROFILARIA AND BRUGIA ANKYRIN
;    TITLE OF INVENTION:  PROTEINS, NUCLEIC ACID MOLECULES, AND
;    TITLE OF INVENTION:  USES THEREOF
;    NUMBER OF SEQUENCES:  85
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Carol Talkington Verser, Ph.D.
;      ADDRESSEE:  Heska Corporation
;      STREET:  1825 Sharp Point Drive
;      CITY:  Fort Collins
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80525
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect for Windows, Version 7.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/041,485
```

; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/847,429
; FILING DATE: 24 APR 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW 5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5503 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-09-031-495-34

Query Match 7.2%; Score 82.6; DB 1; Length 5503;
Best Local Similarity 47.6%; Pred. No. 2.9e-14;
Matches 244; Conservative 0; Mismatches 269; Indels 0; Gaps 0;

Qy 523 actgtctacatctggctctgccaatgggaattcagaagtagtaaaactcctgctggac 582
Db 3815 ACGCCGTTGCATTTAGCTGCTAAGTATGGGAATTGCCGGTCGCGAAATCATTGCTAGAA 3756

Qy 583 agacgatgtcaacttaatgtccttgacaacaaaaagaggacagctctgataaaggccgta 642
Db 3755 CGAGGAACACCGGTTGACATTGAAGGCAAGAATCAGGTAACACCTCTGCATGTAGCGGCA 3696

Qy 643 caatgccagggaagatgaatgtgcgttaatgttgctggaacatggcactgatccaaatatt 702
Db 3695 CATTACAATAACGACAAGGTAGCATTGTTACTTCTAGAAAATGGTGCTTCTGCACATGCC 3636

Qy 703 ccagatgagtatggaaataccactctgcactacgctatctataatgaagataaattaatg 762
Db 3635 GCTGGCAAGAATGGGTACACTCCTTTACATATTGCCGCGAAGAAGAATCAGATGGATATT 3576

Qy 763 gccaaagcactgctcttatatggtgctgatatcgaatcaaaaaacaagcatggcctcaca 822
Db 3575 GCTAGCACTCTCCTTCATTATAAGGCAAATGCGAATGCTGAAAGCAAAGCTGGCTTTACA 3516

Qy 823 ccactgttacttggtgtacatgagcaaaacagcaagtcgtgaaatttttaatcaagaaa 882
Db 3515 CCACTTCATCTTGCCGCCCAGGAGGGCATCGCGAAATGGCTGGGTTATTAATTGAAAAT 3456

Qy 883 aaagcgaatttaaatgcactggatagatatggaaggactgctctcatacttgctgtatgt 942
Db 3455 GGAGCAAAAGTTGGAGCTCAGGCAAGGAATGGCTTGACACCAATGCATTTATGTGCACAG 3396

Qy 943 tgtggatcagcaagtatagtcagccttctacttgagcaaaatattgatgtatcttctcaa 1002
Db 3395 GAGGATCGTGTGAGCGTAGCAGAAGAACTAGTGAAAGAAAACGCAGCCATTGATCCCAAA 3336

Qy 1003 gatctatctggacagcggccagagagtatgct 1035
Db 3335 ACGAAAGCAGGATATACGCCGTTACACGTTGCT 3303

RESULT 14
US-08-847-429A-32
; Sequence 32, Application US/08847429A
; Patent No. 5827692
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/847,429A
; FILING DATE: 24 APR 1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW 5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5503 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 51..5285
US-08-847-429A-32

Query Match 7.2%; Score 82.6; DB 1; Length 5503;
Best Local Similarity 47.6%; Pred. No. 2.9e-14;
Matches 244; Conservative 0; Mismatches 269; Indels 0; Gaps 0;

Qy 523 actgtctacatctggctctgccaatgggaattcagaagtagtaaaactcctgctggac 582
Db 1689 ACGCCGTTGCATTTAGCTGCTAAGTATGGGAATTGCCGGTCGCGAAATCATTGCTAGAA 1748

Qy 583 agacgatgtcaacttaatgtccttgacaacaaaaagaggacagctctgataaaggccgta 642
Db 1749 CGAGGAACACCGGTTGACATTGAAGGCAAGAATCAGGTAACACCTCTGCATGTAGCGGCA 1808

Qy 643 caatgccagggaagatgaatgtgcgttaatgttgctggaacatggcactgatccaaatatt 702
Db 1809 CATTACAATAACGACAAGGTAGCATTGTTACTTCTAGAAAATGGTGCTTCTGCACATGCC 1868

Qy 703 ccagatgagtatggaaataccactctgcactacgctatctataatgaagataaattaatg 762
Db 1869 GCTGGCAAGAATGGGTACACTCCTTTACATATTGCCGCGAAGAAGAATCAGATGGATATT 1928

Qy 763 gccaaagcactgctcttatatggtgctgatatcgaatcaaaaaacaagcatggcctcaca 822
Db 1929 GCTAGCACTCTCCTTCATTATAAGGCAAATGCGAATGCTGAAAGCAAAGCTGGCTTTACA 1988

Qy 823 ccactgttacttggtgtacatgagcaaaacagcaagtcgtgaaatttttaatcaagaaa 882
Db 1989 CCACTTCATCTTGCCGCCCAGGAGGGCATCGCGAAATGGCTGGGTTATTAATTGAAAAT 2048

Qy 883 aaagcgaatttaaatgcactggatagatatggaaggactgctctcatacttgctgtatgt 942
Db 2049 GGAGCAAAAGTTGGAGCTCAGGCAAGGAATGGCTTGACACCAATGCATTTATGTGCACAG 2108

Qy 943 tgtggatcagcaagtatagtcagccttctacttgagcaaaatattgatgtatcttctcaa 1002
Db 2109 GAGGATCGTGTGAGCGTAGCAGAAGAACTAGTGAAAGAAAACGCAGCCATTGATCCCAAA 2168

```
Qy    1003 gatctatctggacagacggccagagagtatgct 1035
                |  | |||  | ||||  |     |  |    ||||
Db    2169 ACGAAAGCAGGATATACGCCGTTACATGTTGCT 2201


RESULT 15
US-08-847-429A-34/c
; Sequence 34, Application US/08847429A
; Patent No. 5827692
;  GENERAL INFORMATION:
;    APPLICANT: Tang, Liang
;    APPLICANT: Blehm, E. Scot
;    TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
;    TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
;    TITLE OF INVENTION: USES THEREOF
;    NUMBER OF SEQUENCES: 85
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE: Carol Talkington Verser, Ph.D.
;      ADDRESSEE: Heska Corporation
;      STREET: 1825 Sharp Point Drive
;      CITY: Fort Collins
;      STATE: Colorado
;      COUNTRY: USA
;      ZIP: 80525
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE: Floppy disk
;      COMPUTER: IBM PC compatible
;      OPERATING SYSTEM: Windows 95
;      SOFTWARE: WordPerfect for Windows, Version 7.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER: US/08/847,429A
;      FILING DATE: 24-APR-1997
;      CLASSIFICATION: 435
;    ATTORNEY/AGENT INFORMATION:
;      NAME: Verser, Carol Talkington
;      REGISTRATION NUMBER: 37,459
;      REFERENCE/DOCKET NUMBER: HW-5
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE: 970/493-7272
;      TELEFAX: 970/484-9505
;  INFORMATION FOR SEQ ID NO: 34:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH: 5503 nucleotides
;      TYPE: nucleic acid
;      STRANDEDNESS: single
;      TOPOLOGY: linear
;    MOLECULE TYPE: cDNA
US-08-847-429A-34


  Query Match             7.2%;  Score 82.6;  DB 1;  Length 5503;
  Best Local Similarity   47.6%;  Pred. No. 2.9e-14;
  Matches  244;  Conservative    0;  Mismatches  269;  Indels    0;  Gaps    0;

Qy     523 actgctctacatctggcctctgccaatgggaattcagaagtagtaaaactcctgctggac 582
Db    3815 ACGCCGTTGCATTTAGCTGCTAAGTATGGCAATTTGCCGGTCGCGAAATCATTGCTAGAA 3756

Qy     583 agacgatgtcaacttaatgtccttgacaacaaaaagaggacagctctgataaaggccgta 642
Db    3755 CGAGGAACACCGGTTGACATTGAAGGCAAGAATCAGGTAACACCTCTGCATGTAGCGGCA 3696

Qy     643 caatgccaggaagatgaatgtgcgttaatgttgctggaacatggcactgatccaaatatt 702
Db    3695 CATTACAATAACGACAAGGTAGCATTGTTACTTCTAGAAAATGGTGCTTCTGCACATGCC 3636

Qy     703 gcagatggatatggaaatacgactctgcatatagctatctataatggagataattaatg 762
Db    3635 GCTGCCAAGAATGGGTACACTCCTTTACATATTGCCGCGAAGAAGAATCAGATGGATATT 3576

Qy     763 gccaaagcactgctcttatatggtgctgatatcgaatcaaaaaacaagcatggcctcaca 822

Db    3575 GCTAGCACTCTCCTTCATTATAAGGCAAATGCGAATGCTGAAAGCAAAGCTGGCTTTACA 3516

Qy     823 ccactgttacttggtgtacatgagcaaaaacagcaagtcgtgaaattttaatcaagaaa 882
Db    3515 CCACTTCATCTTGCCGCCCAGGAGGGCCATGCGAAATGGCTGCGTTATTAATTGAAAAT 3456

Qy     883 aaagcaatttaaatgcactggatagatatggaaggactgctctcatacttgctgtatgt 942
Db    3455 GGAGCAAAAGTTGGAGCTCAGGCAAGGAATGGCTTGACACCAATGCATTTATGTGCACAG 3396

Qy     943 tgtggatcagcaagtatagtcagccttctacttgagcaaaatattgatgtatcttctcaa 1002
Db    3395 GAGGATCGTGTGAGCGTAGCAGAAGAACTAGTGAAAGAAAACGCAGCCATTGATCCCAAA 3336

Qy    1003 gatctatctggacagacggccagagagtatgct 1035
Db    3335 ACGAAAGCAGGATATACGCCGTTACATGTTGCT 3303


Search completed: March 9, 2002, 02:21:27
Job time: 6251 sec
```